Christoph Fischer

Establishment of the Luciferase Assay for the identification of direct microRNA/messenger RNA interactions in human cells

Diploma Thesis

Institute for Genomics and Bioinformatics
Graz University of Technology
Petersgasse 14, 8010 Graz

Supervisor:

Univ. Ass. Dipl. Chem. Dr Marcel Scheideler

Evaluator:

Univ. Ass. Dipl. Chem. Dr Marcel Scheideler

Graz, October, 2010

STATUTORY DECLARATION

I declare that I have authored this thesis independently, that I have not used other than the declared sources / resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

…… ……..
date (signature)

Abstract

Background:

MicroRNAs, a class of regulatory ribonucleic acids, intervene in the fate of protein coding messenger RNAs by binding specifically to them, thereby positioning associated proteins spatially such that they can interact with the protein groups regulating translation and mRNA decay. The effect of interaction ranges from repression or under circumstances enhancement of translation via accelerated decay right up to the immediate destruction of the mRNA strand. In any case the cellular amount of the respective protein is influenced. Present findings assign microRNAs a crucial role in virtually all cellular processes, from proliferation via differentiation and maintenance of the cell identity through to apoptosis. Accordingly, many diseased states (with cancer leading the way) exhibit abnormal microRNA expression profiles. All these circumstances immediately suggest the idea to utilize microRNAs therapeutically.

Objective:

Computer-aided prediction of functional microRNA/mRNA pairs is very demanding with respect to physical-mathematical modeling, due to this fact results may be assessed as rather unreliable so far. Therefore laboratory methods of validation are absolutely necessary to identify functional pairs. The *luciferase assay* is a suitable system for *in vitro* verification of both specific binding and especially functional interaction. The establishment of this method was the primary aim of the present work.

Additionally, a first case of application should proof the operability of the validation system, and support indications from bioinformatic predictions and other experimental methods gained in the specific study. Due to its central role in the gene regulatory network of fat cells and a predicted, strongly conserved microRNA binding site in its mRNA, the nuclear receptor PPARγ (*peroxisome proliferator-activated receptor-γ*) was chosen as first candidate for validation.

Result:

The *luciferase assay* could successfully be established for the 96-well format. The small form factor and optimizations on transfection and cell culture conditions allow high efficiency of the test system with corresponding beneficial economic consequences. The handling was kept as easy as possible and instruction sheets are available to the user.

Validation of the interaction pair PPARγ and miR-27b was positive. Thus, miR-27b was identified as one of the first microRNAs with significant influence in the regulatory network of human fat cells (publication (Karbiener et al. 2009)).

Zusammenfassung

Hintergrund:

MicroRNAs, eine Klasse regulativer Ribonukleinsäuren, greifen in das Schicksal von proteinkodierenden messenger RNAs ein, indem sie spezifisch an sie binden und dadurch assozierte Proteine räumlich derart positionieren, dass sie mit Proteingruppen, welche die Translation bzw. den Abbau von mRNAs durchführen, interagieren können. Die Wirkungen dieser Interaktion reichen dabei von Hemmung oder unter Umständen Verstärkung der Translation über beschleunigten Abbau bis hin zur direkten Zerstörung des mRNA-Stranges. In jedem Fall wird Einfluss auf die zellulär vorliegende Menge des jeweiligen Proteins genommen. Bisherige Ergebnisse weisen den microRNAs eine entscheidende Rolle in praktisch allen zellulären Prozessen zu, von Zellteilung über Differenzierung und Erhalt der Zellidentität bis hin zum geregelten Zelltod. Entsprechenderweise zeigen viele krankhafte Zustände (allen voran Krebs) auch abnormale microRNA Expressionsprofile. Diese Umstände legen den Gedanken nahe, microRNAs therapeutisch einzusetzen.

Ziel:

Computergestütztes Vorhersagen von funktionellen microRNA/mRNA Paaren ist hinsichtlich der physikalisch-mathematischen Modellbildung sehr anspruchsvoll, diesem Umstand geschuldet sind die Ergebnisse bisweilen als eher unzuverlässig zu bewerten. Deshalb sind labortechnische Validierungsverfahren absolut notwendig um funktionelle Paare zu identifizieren. Der *Luciferase Assay* ist ein geeignetes Verfahren, um spezifische Bindung und vor allem funktionelle Interaktion *in vitro* nachzuweisen. Die vollständige technische Etablierung dieser Methode war primäres Ziel der vorliegenden Arbeit.

Außerdem sollte ein erster Anwendungsfall die Funktionsfähigkeit des Validierungssystems zeigen und Hinweise aus bioinformatischen Vorhersagen und anderen experimentellen Methoden stützen, die aus der konkreten Studie hervorgingen. Wegen seiner zentralen Rolle im Genregulationsnetzwerk von Fettzellen und einer vorhergesagten, stark konservierten microRNA Bindungsstelle in seiner mRNA war der Kernrezeptor PPARγ (*peroxisome proliferator-activated receptor-γ*) der erste Testkandidat für eine Validierung.

Ergebnisse:

Der *Luciferase Assay* konnte erfolgreich für das 96-well Format etabliert werden. Der kleine Formfaktor und Optimierungen der Transfektions- und Zellkulturbedingungen ermöglichen eine hohe Effizienz des Testsystems mit entsprechend günstigen ökonomischen Folgen. Die Handhabung wurde so einfach wie möglich gehalten und entsprechende Arbeitsanleitungen liegen dem Anwender vor.

Die Validierung des Interaktionspaares PPARγ und miR-27b war positiv. miR-27b wurde damit als eine der ersten microRNAs mit wesentlichem Einfluss im Regulationsnetz von humanen Fettzellen identifiziert (Publikation (Karbiener et al. 2009)).

Contents

1. Introduction

The current core theme at the *Institute for Genomics and Bioinformatics* (IGB) is *adipogenesis*, the development of precursor cells to mature fat cells. Although the established *luciferase assay* is universally usable for microRNA studies – and with minor changes also for promoter and transcription factor studies, respectively – the context of its (at present) real application shall be pointed out. Therefore the first three sections give a rather detailed but compact summary on the current state of research regarding obesity, metabolism, fat cell development, and molecular factors likely involved in adipose tissue-related regulatory networks and diseases.

1.1 Obesity – a modern world disease

Improved standards of living and increased access to high-calorie diet combined with decreased physical activity and inappropriate dietary patterns have led to a strong rise in obesity and diet-related diseases worldwide. Clearly, the main determinant of obesity epidemic is energy imbalance: more calories consumed than expended. According to estimations (based on body mass index; see below) of the *World Health Organization* (WHO) about 1 Billion adults worldwide are overweight, and at least 300 Million are classified as obese, with both numbers seriously increasing over the last few decades. A similar trend can be seen with children and adolescents. 22 Million children under the age of 5 are estimated to be overweight worldwide. Startling, as apparently obesity is a major factor in the development of several medical conditions such as type-2 diabetes, cardiovascular diseases, hypertension, stroke, atherosclerosis, Alzheimer's disease and certain cancers. These conditions might be caused by obesity-induced *insulin* resistance and especially the fact that adipose tissue serves not only as simple energy depot but also as endocrine organ (indeed the body's largest) secreting hormones and cytokines that effect the function of cells all over the body (WHO 2010).

In most animals excess energy storage occurs in the form of *triglycerides* (*fat*) accumulated in a mesodermal tissue called *white adipose tissue* (WAT). The distribution of WAT varies considerably between species and to some degree also between individuals of the same. In general, in humans WAT is dispersed throughout the body with major intra-abdominal depots (*visceral fat*; coating the inner organs), as well as *subcutaneous* storage places (especially in the buttocks, thighs, and abdomen). In addition, WAT can be found in many other areas including the face and extremities, where it may primarily act as thermal insulator or surface modeler or has its main function in the provision of mechanic protection and support. The most accurate approaches available for measuring a body's fat mass and distribution are *magnetic resonance imaging* (MRI; expensive), and *dual-energy x-ray absorptiometry* (DXA; expensive and radiation exposure). But for clinical purposes simple surrogate measurements like *body mass index* (BMI) and *waist-to-hip ratio* (WHR) are mostly sufficient and favored (Gesta et al. 2007).

BMI is only a crude measure of fat mass but the most useful one to determine overweight and obesity on population-level as it is the same for both sexes and for all ages of adults (under the age of 18 a different age-dependent cut-off table is applied). It is calculated as a person's weight (in kilograms) divided by the square of the height (in meters), with an index of ≥ 25 [kg/m²] being considered as overweight and ≥ 30 as obese – though, ethnic specific differences in body

composition cause discussions on the need of adjustments of cut-off values for different populations with respect to health aspects. Although BMI values between 18.5 and < 25 are defined as normal range, there is evidence that risk of chronic disease in populations increases progressively from a BMI of 21 (WHO fact sheet No311 2006). The WHR (ratio of the circumference of the waist to that of the hips) on the other hand is a simple method that allows for determination of the distribution of adipose tissue. Twin and population studies have revealed that BMI and WHR are heritable traits (Nelson 2000), and thus genetics plays an important role in both amount and distribution of accumulated WAT – clearly, genes only determine a person's susceptibility to fat accumulation. Increases in fat mass are based on increased amounts of intracellular lipids and greater adipocyte size (hypertrophy) and an increased number of cells (hyperplasia). Hypertrophy has been considered the sole route of adipose tissue enlargement in adults (Hirsch & Batchelor 1976), but by now it is known that also hyperplasia contributes to the increased WAT masses in adult-onset obesity, whereat the dominant mechanism may vary with location of the fat depot. In vitro studies suggest that factors released by hypertrophied adipocytes, such as *tumor necrosis factor-α* (TNF-α) and *insulin-like growth factor-1* (IGF-1), stimulate hyperplasia in a paracrine fashion (affecting only nearby cells). According to animal studies, hyperplasia is caused by an increase in numbers of preadipocytes in the first place (paracrine) and subsequent differentiation of these precursor cells into mature (adipokine-secreting) adipocytes. The regulating factors of differentiation are not fully understood, but circulating *insulin* and *glucocorticoid* concentrations appear to stimulate this course of events.

Fat distribution or rather the size of certain WAT depots consisting of different fat cell (*adipocyte*) populations seems to play a major role in metabolism as increased visceral/intra-abdominal fat (typical male pattern; high WHR) is related to a higher risk of metabolic disease, whereas increased subcutaneous fat in the thighs and hips (typical female pattern; low WHR) exerts little or no risk (Kissebah & Krakower 1994). A possible and likely explanation for this location dependency may be that at least some different WAT depots are made up of white adipocyte populations with different properties. Indeed, expression profiling revealed significant differences in expression of hundreds of genes between cells from different depots (Vidal 2001, Vohl et al. 2004) and even within a single fat depot (Tchkonia et al. 2005). The probably intrinsic variation (that also persists in culture) may cause different metabolic activity and also altered reactions on extrinsic factors or patterns of secreted hormones.

These different types of white adipocytes are thought to result from separate developmental lineages (see section **1.3** for some recent details) (Gesta et al. 2007). Interestingly, aging is associated with increases in visceral fat with concomitant loss of subcutaneous depots (Kuk et al. 2009).

Due to strikingly different tasks, the adipose organ in mammals can be functionally classified into two main categories of fat cells. Besides the white adipocytes forming the WAT and primarily responsible for energy storage, induced release of free fatty acids, and hormonal regulation of energy homeostasis (with apparent heterogeneity also introducing functional diversity, though), the so called *brown adipose tissue* (BAT) formed by *brown adipocytes* is important for basal and especially inducible energy dissipation as heat during cold- and diet-induced thermogenesis. Also in white and brown adipocytes many genes are differentially expressed, but in the essential points both lineages show similarities as they are marked by *insulin*-regulated *glucose* uptake, and the presence of PPARγ (see below) and other markers of terminal differentiation (Rosen & MacDougald 2006). But, in addition WAT is characterized by the presence of the proteohormone *leptin*, while BAT can be distinguished by the expression of *uncoupling protein-1* (UCP-1).

UCP-1 (also referred to as *thermogenin*) is a mitochondrial proton carrier (a channel protein) in the inner mitochondrial membrane of brown adipocytes. It is used to generate heat by non-shivering thermogenesis per uncoupling the phosphorylation reaction taking place in ATPsynthases from the electron transport chain (for details see next section). Though, experimental results from UCP1-deficient mice suggest the presence of alternative, apparently less efficient thermogenic mechanisms (consuming more fuel and oxygen for generation of a certain amount of heat), taking place in brown adipocytes developing in WAT depots upon cold exposure. Interestingly, a double knockout of UPC-1 and the adipokine *leptin* renders mice unable to adapt to temperatures below 12°C, unless they are administered either *leptin* or *thyroid hormone* (Ukropec, Anunciado, Ravussin, Hulver et al. 2006, Ukropec, Anunciado, Ravussin & Kozak 2006) (the latter apparently being the end-effector as the hormone cascade goes like *thyrotopin-releasing hormone*, TRH (origin: hypothalamus; brain) → *thyroid stimulating hormone*, TSH (origin: pituitary; brain) → *thyroid hormone* (origin: thyroid; throat); where the production of TRH is controlled by *leptin* directly at the transcriptional level (Shibusawa et al. 2008)). A more recently published study by Ribeiro et al. reports the connection of *thyroid hormone receptor-β isoform* (TR β) with adaptive thermogenesis, though, the group links the defective adaptive thermogenesis in TR β mutant mice to the expression of UCP-1 being reduced thereby (Ribeiro et al. 2010) – what might be only half the story.

Both, BAT and WAT are strongly innervated by the autonomic nervous system (ANS) and underlie extensive central nervous regulation of metabolism, thermogenesis, and secretory activity. WAT and the liver play the central roles in fuel storage and release, whereat both organs respond to altered energy availability with a set of homeostatic responses mediated by humoral factors and ANS signaling. For instance, activation of hepatic sympathetic innervation increases *glucose* output and modulates fatty acid transport, while parasympathetic activity decreases *glucose* output and increases carbohydrate storage. Similarly, in WAT the activation of sympathetic innervation induces lipolysis and alters *glucose* uptake. In general it can be said that parasympathetic activity promotes fuel storage, whereas sympathetic activity increases the amount of fuel available for immediate use (Stanley et al. 2010). Thermogenesis in BAT on the other hand is apparently controlled by *norepinephrine* (*noradrenalin*) released from cells of the sympathetic nervous system, and interacting mainly with β-adrenergic receptors to stimulate thermogenesis (and also white to brown transdifferentiation, see below) (Pénicaud 2010, Huang et al. 2010, Bamshad et al. 1999).

Morphologically, BAT can be distinguished from WAT by multilocular lipid droplets (unilocular in WAT), richer vascularization (necessary to satisfy the high oxygen needs of brown adipocytes) and high mitochondria density (causing the color). BAT has also a different developmental pattern. Where WAT depots in humans are set up beginning mid gestation and gradually increase in volume throughout life where at birth both visceral and subcutaneous depots are apparent, BAT typically emerges earlier during fetal development and is at its maximal size (relative to body weight) at birth, when nonshivering thermogenesis is needed the most, then declines with aging (Cannon & Needergaard 2004). BAT has been thought to be virtually absent in human adults, but tumor metastasis screening with co-registered *fluorodeoxyglucose positron emission tomography* and *computed tomography* (FDG PET/CT) has accidentally revealed the main localizations of this tissue in the supraclavicular (above the collar bone) and the neck regions, and in addition some smaller depots paravertebral (around the backbone), paraaortic (around the aorta, the largest artery in the body originating directly from the left ventricle of the heart and extending down to the abdomen where it branches into smaller arteries, moreover, in the aortic arch – the aorta's first bend close to the heart – the carotid arteries supplying the brain branch off), mediastinal (between the lungs) and suprarenal

(above the kidney). With respect to heat distribution and supply of heat for the most important organs (all depots), and also heat dissipation into the environment (the main depots) these locations are well selected. Brown adipocytes are also found interspersed in WAT depots after cold acclimatization (Saverio Cinti 2006, Barbatelli et al. 2010). The exact origin of these cells is still unclear, but they seem to have a mixed morphology and different populations of mitochondria (classic "white" and "brown" ones) suggesting a direct white to brown transformation process, where this transdifferentiation apparently depends on *beta(3)-adenoceptor* activation (Barbatelli et al. 2010).

However, the prevalence of active BAT is unknown so far, and it may be that only a fraction of adult humans possess BAT with metabolic significance - available results are inconclusive showing proportions ranging from 2.5% (638 cases in total; (Hany et al. 2002) over 4% (32 female, 17 male patients; age range 12-77 years; (Cohade et al. 2003)) to 45% (33 cases in total; (Rousseau et al. 2006)), and up to 96% under explicit cold exposure (16°C; 24 cases in total; all young men aged 18 to 32; (van Marken Lichtenbelt et al. 2009)). Indeed the results ranging from 2.5 to 45% (without explicit cold-induction) may just reflect bad experimental control as FDG PET procedures are typically associated with slight cold stress (Nedergaard et al. 2007) what may cause different proportions of cold-induced activity even if the same group of patients would be tested in the different clinical environments.

Assumable interindividual differences in amount and (especially diet-induced) activity of BAT could be an explanatory contributory factor for different susceptibility to obesity in humans. Likewise age-related decrease of both BAT mass and activity may be linked to the general trend of increasing WAT mass over lifetime. Indeed, there are estimations that as little as 50 g BAT could account for about 20% of a body's daily energy expenditure, if maximally stimulated (Gesta et al. 2007). Although, some statistical analyses suggest correlations between BMI, age, and sex (independently) with BAT activity (Pfannenberg et al. 2010), (van Marken Lichtenbelt et al. 2009), also these potential relationships need further investigation due to inconsistency (e.g. (Cohade et al. 2003)). There seem to be gender specific differences with higher BAT mass (32 ± 5 g vs. 18 ± 4 g; $p \leq 0.0006$) and activity (1.59 ± 0.1 vs. 1.02 ± 0.1; $p \leq 0.0006$) in women under controlled thermoneutral conditions (Pfannenberg et al. 2010). Though, in the conducted study these differences only became apparent between female and male subjects in the upper two age tertiles, what has to be seen against the background of a seemingly much stronger decline of BAT mass and activity with increasing age in men. But, briefly and generally speaking, both BAT mass and BAT activity may be elevated in women and in younger people.

Anyhow, since brown adipocytes undoubtedly possess a very high capacity to dissipate excess energy, thereby preventing energy storage in WAT, therapeutically increased amounts of BAT are an obvious approach for treating obesity.

1.2 Metabolism – in essence

Cells need to do work in order to preserve the complex structures of life, since organization is instable by nature. The energy necessary for these operating procedures is provided by nutrients in the diet and in most ecosystems by sun light in the end, respectively.

In humans, an extensive network of enzymatic reactions made up of several thousand reactions carries out the transformation and syntheses of molecules, and also the conversion of energy in the nutrients to adenosine triphosphate (ATP), ***the*** carrier of Gibbs energy (ΔG) in the body. Altogether these chemical reactions are referred to as *metabolism*, and can be categorized into *catabolism* (reactions making energy available) and *anabolism* (reactions using this energy for biosynthesis). A reaction being thermodynamically unfavorable (because the change in free energy is not negative) can be coupled to a favorable one by specific enzymatic reactions (e.g. phosphorylation), whereat this energetically favorable reaction is the hydrolysis of ATP in most cases. ATP hydrolysis typically changes the equilibrium of a coupled reaction by a factor of ~ 10^8 to the (otherwise unlikely) product side. This strong shift is driven by the instability ("high energy content") of the phosphoanhydride bonds in ATP molecules (under cellular environment conditions) and especially by the high ratio of ATP to ADP concentrations (far from equilibrium) typically maintained by the ATP generating system. ATP not only drives chemical synthesis but is also used for signal transduction and other energy demanding processes, such as mechanical work (e.g. contraction of muscle cells) and active transport processes (e.g. generation of concentration gradients over membranes).

In principle, all reaction network activity is centered on a small number of activated carriers like ATP (providing approx. -50 kJ/mol *Gibbs* energy and phosphate group-transfer potential, respectively), NADH, $FADH_2$ (both providing redox potential), *pyruvate* (main feeder of the citric acid cycle providing the first intermediate; see below), and *acetyl-CoA* (main feeder of the citric acid cycle providing activated acyl groups and carbon, respectively) forming the core of metabolism, and of course O_2 (as highly electronegative electron acceptor) and H_2O (as H and electron donor, respectively).

Excess energy and thus storage in fat depots can be caused by all nutrients as the processing pathways of *glucose* (carbohydrates, polysaccharides), fatty acids (triglycerides), and amino acids (peptides, proteins) are linked by *acetyl-CoA* (*acetyl coenzyme A*; activated acetic acid) and the *citric acid cycle* (*tricarboxylic acid cycle*, TCA cycle) (**Fig. I1**).

In their catabolic pathways amino acids enter the TCA cycle species-specifically either indirectly by being decomposed to *pyruvate* or *acteyl-CoA* (see **Fig. A1**, Appendix), or directly after being processed to *oxaloacetate*, *succinyl-CoA*, *fumarate*, or *α-ketoglutarate* – all intermediates of the TCA cycle (for molecules forming the TCA cycle and more details on fluxes see **Fig. A1**, Appendix). Beside some glycolytic intermediates and *pyruvate*, *α-ketoglutarate* and *oxaloacetate* are also the basis for amino acid synthesis. If not used in synthesis or energy production, carbon skeletons of amino acids are generally conserved in the form of *glucose* (per *gluconeogenesis*) or as fatty acids (per *fatty acid synthesis*). However, with respect to eating habits and energy turnover fatty acid and carbohydrate metabolism are of primary importance as compared with protein metabolism. Related catabolic processes (*glycolysis* and *β-oxidation*) and the reverse anabolic processes (*gluconeogenesis* and *fatty acid synthesis*) balance the carbohydrate (*glucose*/*glycogen* – the storage form, approximately 100 to 150 g in the liver and 300 to 400 g in total) and lipid (fatty acids/triglycerides – the storage form)

amounts and transform mainly by the metabolic intermediate *acetyl-CoA* and the TCA cycle one species into the other according to offers and needs, respectively – where it is to mention that *glucose cannot* be synthesized solely from fatty acids, as β-oxidation does not provide any TCA cycle intermediates besides the fuel *acetyl-CoA*. But glucogenic amino acids may provide the necessary chemical groups and intermediates, respectively (see **Fig. A1**, Appendix).

Among others, especially two pancreatic hormones are responsible for the regulation of blood *glucose* levels and related metabolic processes, *insulin* and *glucagon*. For instance, in case of excess carbohydrates available as energy source, some *glucose* is completely oxidized to CO_2 and H_2O (*glycolysis*, TCA cycle, and *respiratory chain*), some of it is stored as *glycogen* (a branched glucose polymer being the animal analogue of starch; stored in the liver, muscles and in minor amounts in several other tissues) per *glycogenesis* (in the liver activated by resting periods due to the *lactate-glucose cycle*, by high *glucose* levels in the blood that are autonomously sensed, and also by *insulin*; only liver stores are used to stabilize blood *glucose* levels), and some of it is stored by using conversion to *acetyl-CoA* and subsequent creation of fatty acids (followed by esterification with *glycerol* to form triglycerides, together referred to as *lipogenesis*; also activated by *insulin*) and *cholesterol* (both via TCA cycle intermediate *citrate*). When low amounts of *glucose* are available (due to starvation, low uptake of carbohydrates with the diet but also regularly in the time during meals), *glycogen* reserves in the liver are consumed (where the break down into *glucose* is referred to as *glycogenolysis*; among others stimulated by *glucagon* in response to low *glucose* levels) and fatty acids are set free from triglycerides (referred to as *lipolysis*; among others activated by the hormones *glucagon*, *adrenalin*, *noradrenalin*, *growth hormone*, and *cortisol*) and broken down into *acetyl-CoA*, which is fed into the TCA cycle and ends up in *gluconeogenesis* (via *malate* and *oxaloacetate*) and ATP production (via $NADH/H^+$ and $FADH_2$, that are fed into the *respiratory chain*), respectively. High levels of *acetyl-CoA* either caused by high feeding rates from glycolysis, protein catabolism, and fatty acid oxidation or low amounts (or activity) of TCA cycle intermediates (e.g. caused by low *glucose* and thus *pyruvate*/*oxaloacetate* levels with subsequent depletion of the TCA cycle intermediates due to gluconeogenesis) overcharge the TCA cycle and increase the production of ketone bodies (*ketogenesis*). On *glucose* shortage ketone bodies are used as vital energy source in the brain and to a lesser extent in the heart (which generally uses mainly fatty acids, what the brain is not capable of) by reconversion to *acetyl-CoA*.

In general, the production of ATP molecules and likewise of all other synthesis products and intermediates mentioned or depicted (and also of those omitted) in **Fig. I1** and **Fig. A1** is *tightly regulated* by feedback effects (e.g. *hexokinase* in glycolysis; directly inhibited by its product *glucose-6-phosphate*), allosteric mechanisms (e.g. *phosphofructokinase* (PFK) in glycolysis; allosterically inhibited by ATP and activated by ADP), and by substrate concentration dependencies of all the (omitted) enzymes within the different pathways. Key control points arise in enzymatic reactions being so energetically favorable that they are effectively irreversible under physiological conditions (like PFK in glycolysis). The TCA cycle is mainly regulated by the availability of key substrates such as NAD^+, NADH, P_i, Ca^{2+} and the adenosine pool ATP, ADP and AMP. Moreover, *citrate*, the first product in the cycle (**Fig. A1**), is an immediate feedback inhibitor of its own productive enzyme *citrate synthase* and also of PFK, providing the link to glycolysis. Oxidative phosphorylation is mainly controlled by the reaction catalyzed by *cytochrome c oxidase* (the complex IV) – the oxidation of *cytochrome c* and reduction of final electron acceptor O_2 to H_2O – as a function of the ratios NADH to NAD^+ (upstream energy output) and ATP to ADP+Pi (energy load). So, in case of high ATP levels and virtually absent ADP molecules, the proton gradient cannot be reduced by ATPsynthases what in turn

stems the electron transport chain causing an accumulation of NADH and depletion of NAD^+ falling back on the TCA cycle and glycolysis also in a decelerating manner. Accordingly, in the opposite case of high ADP levels this mechanism accelerates the respiratory chain and all upstream processes. This system behavior, referred to as *respiratory control*, is of fundamental importance for ATP management, and also for effects of uncoupled states (see below).

The amount of ATP in a human body is typically about 100 g, but the ATP-turnover is extremely high. For instance, in normally active muscle cells the total ATP pool is typically consumed and regenerated within a minute what means that *per second* 10 Million ATP molecules *per cell* are consumed and regenerated. Accordingly, a human in resting state spends an ATP mass of about half the body weight in 24 hours (turnover of approximately the body weight), and under intensive work load this value can increase to about 0.5 kg of ATP per minute (e.g. 60 kg of ATP for driving a 2-hour run).

(Stryer 2003, Campbell 2003)

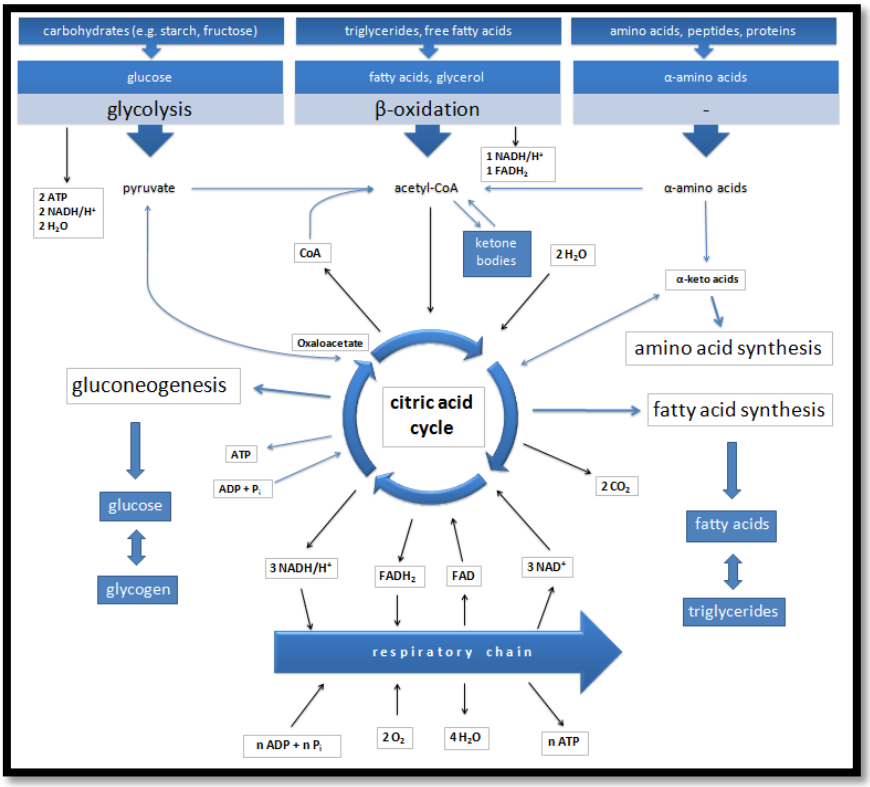


Fig. I1 Overview over central metabolic processes and some important involved molecules (blue arrows represent enzymatic reactions). The sketched reaction sequence from nutrients (e.g. carbohydrates and triglycerides) to the final products CO_2 and H_2O is driven by diminishing *Gibbs* energy in each step. ATP synthesis is coupled to the oxidation of carbon compounds, either directly (within the frame of substrate-level phosphorylation e.g. in glycolysis and the TCA cycle) or, and that is the pathway producing most ATP molecules, via formation of proton gradients in the respiratory chain. The energy recovery from nutrients can be separated into three stages. In the first step, larger molecules such as starch, fat and proteins are broken down into the respective building blocks *glucose*, fatty acids and amino acids by a multitude of enzymes (amylases, proteases etc.) whereby no available energy is released. In the second step these building blocks are further decomposed to smaller, simpler units such as *pyruvate* and *acetyl-CoA*, the latter being a carrier of activated acyl groups and one of the central molecules in metabolism. In the third stage, involving the citric acid cycle (TCA cycle) and the *respiratory chain* (or *oxidative phosphorylation*; comprising the *electron transport chain*), the nutrients are fully oxidized to CO_2 and H_2O. In the TCA cycle carbon oxidation reactions transfer the energy

of *acetyl-CoA* as 'energized' electrons to NAD^+ (nicotinamide adenine dinucleotide; $NAD^+ \rightarrow NADH$), and FAD (flavin adenine dinucleotide; $FAD \rightarrow FADH_2$), which provide redox potential for the electron transport chain. In general, redox reactions release energy by transferring electrons to atoms being more electronegative, that is binding the electrons more closely and thus with lower potential energy (mostly this "electron transfer" is actually a transfer of a hydrogen atom – an electron bound to a proton). The stepwise transfer of electrons from NADH and $FADH_2$ (regenerating the carriers: $NADH \rightarrow NAD^+$; $FADH_2 \rightarrow FAD$) over 4 enzyme complexes (complex I: *NADH-coenzyme Q oxidoreductase*, entry point for NADH; complex II: *Succinate-Q oxidoreductase*, entry point for $FADH_2$ and also part of the TCA cycle; complex III: *Q-cytochrome c oxidoreductase*; complex IV: *cytochrome c oxidase,* due to occasionally occurring release of intermediates mainly responsible for the much-talked about *reactive oxygen species*) in the inner mitochondrial membrane to the final electron acceptor O_2 (being one of the most electronegative atoms) powers the pumping of protons out of the mitochondrial matrix into the mitochondrial intermembrane space. This proton gradient over the inner mitochondrial membrane is the driving force (called *proton motive force;* PMF) for motor proteins (*ATP synthases; complex V*), which produce ATP from ADP (adenosine diphosphate) and P_i (inorganic phosphate) when the protons reduce their electrochemical gradient by running through the synthases back into the matrix (*phosphorylation*).

Why makes a little too much fat a little more fat than a little too much sugar?

Organic molecules with a lot of hydrogen (H) atoms are usually excellent fuels, since the chemical bonds involving H atoms are commonly a source of electrons with high potential energy (see text **Fig. I1**). The decomposition of fatty acids (carbon-hydrogen chains) is a reaction sequence providing a good deal of available energy to the body. For instance, the complete decomposition of a single *palmitate* molecule ($C_{16}H_{32}O_2$; 256.43 $g \cdot mol^{-1}$; a saturated 16 carbon fatty acid; the end-product in fatty acid synthesis and the most common saturated fatty acid in animal and vegetable fats and oils) by β-oxidation, citric acid cycle (1 *acetyl-CoA* per 2 carbons) and respiratory chain yields approximately 100 (98 to 106) molecules of ATP, whereas the complete processing of a single *glucose* molecule ($C_6H_{12}O_6$; 180.16 $g \cdot mol^{-1}$) by glycolysis, citric acid cycle (maximal 2 *acetyl-CoA* per *glucose* molecule) and respiratory chain gives about 30 (30 to 32) molecules of ATP (indeed, the outcome of oxidative phosphorylation is uncertain as the stoichiometry of proton pumps, ATPsynthases and the transport of metabolites are neither integer nor necessarily constant in each case, moreover, the proton gradient also serves for other processes besides driving ATPsynthases such as *pyruvate* uptake and there are losses due to proton and electron leak pathways (Jastroch et al. 2010)). An ATP yield ratio of roughly 3 and a weight ratio of about 1.4 give a bit more than two times the energy content per mass unit for fatty acids as compared with *glucose*. Concerning the storage forms, this fact combined with the high hygroscopicity of *glucagon* (what would increase body weight due to water retention) and the hydrophobicity of triglycerides, makes obvious why long-term energy storage is done with fat – but, *glucagon* on the other hand is able to release energy more quickly.

Especially with regard to heat production of brown adipocytes a rough estimate of the efficiency factor of the energy metabolism is of peculiar interest. The following considerations are representatively carried out for *glucose* under chemical standard conditions, where the relevant total reaction has the form:

$$C_6H_{12}O_6 + 6\,O_2 \rightarrow 6\,CO_2 + 6\,H_2O + energy\ (ATP + heat)$$

Complete oxidation of a mole of *glucose* would yield 2870 kJ (ΔG = -2870 kJ/mol; -686 kcal/mol), a mole of ATP stores 30.5 kJ Gibbs energy. As energy metabolism generates (let's say) 30 molecules of

ATP per *glucose* molecule the efficiency factor is roughly 32 % ($\frac{30*30.5}{2870} \approx 0.32$) – the residual energy is mainly dissipated as heat, which is primarily distributed by body fluids (H_2O molecules) and used to sustain the relatively high body temperature – the excess is delivered to the environment by cooling processes (radiation and heat conduction enhanced by regulation of blood vessel diameter near the body surface and sweating; respiration and water balance). In case of the uncoupled state in brown adipocytes the only difference in the outlined metabolic pathway is the change of the last step, where the proton gradient reduction by ATPsynthases is redirected to UCP-1 channels. This redirection corresponds to short-circuiting of the proton circulation over the inner mitochondrial membrane what depletes the ATP pool and raises the ADP level, respectively, and thereby increases the reaction rates of proton motive force generating processes causing a highly increased demand on fuels and especially O_2 – and of course increased production of heat.

(Stryer 2003, Campbell 2003)

1.3 Adipocytes – some facts on the molecular level

Adipose tissue is an important part of the endocrine system as it releases many (by now ~100 known substances) metabolically active proteins known as *adipokines*, such as the hormones *leptin*, *adiponectin* and *resistin* or the cytokines *tumor-necrosis factor-alpha-α* (TNF-α), *Interleukin-6* (IL-6), and *plasminogen activator inhibitor-1* (PAI-1; a key regulatory protein of – among others – tissue fibrinolysis, cell migration, angiogenesis and tissue remodeling (Lijnen 2005)) (David C. W. Lau et al. 2005). These adipokines regulate systemic processes in an autocrine (self regulation of a cell), paracrine (regulation of nearby cells) and endocrine (systemic) fashion, thereby intervening with energy balance, *insulin* sensitivity, blood pressure, angiogenesis, immune response and lipid metabolism. Especially the hormones *leptin* and *adiponectin* seem to be of major importance as elevated *leptin* and decreased *adiponectin* levels are characteristic for the metabolic syndrome and also correlated with high risk of cardiovascular disease (The metabolic syndrome is a cluster of traits that include *hyperinsulinemia* (excess levels of circulating *insulin*), abnormal *glucose* tolerance, obesity, hypertension and a *dyslipidemia* characterized by high triglycerides, low HDL (high density lipoprotein) cholesterol and small-dense LDL (low density lipoprotein) particles leading to atherosclerosis.). *Leptin* acts within the central nervous system and thereby regulates the intake of food by decreasing appetite with increasing levels of the hormone. Moreover, it stimulates energy expenditure and regulates the immune response. Actually, the level of *leptin* correlates positively with fat mass but in the state of obesity the higher concentrations of *leptin* cannot develop the normal effect due to a lack of sensitivity to its action (the mechanism is still to elucidate). *Adiponectin* plays an important role as *insulin*-sensitizing hormone. In contrast to *leptin* and most other adipokines, concentration of *adiponectin* is lowered in obesity, where low-circulating levels of this hormone are associated with *insulin* resistance, diabetes, hypertension, fatty liver disease and ischemic heart disease (Gnacińska et al. 2009), (Stofkova 2009), (Antuna-Puente et al. 2008), (Luo, Jian Liu, B. Hong Chung et al. 2010), (Fernández-Riejos et al. 2010). Indeed, modestly increased levels of circulating *adiponectin* by ectopic overexpression improved *insulin*-sensitivity and normalized *insulin* and *glucose* levels in ob/ob mice (which are *leptin* deficient and morbidly obese), and thus completely rescued the diabetic phenotype. Moreover, these mice displayed increased expression of PPARγ target genes (see below) and a reduction in systemic inflammation and macrophage infiltration in adipose tissue (Ja-Young Kim et al. 2007).

Adipocyte origins

So far, little is known about the developmental origins of adipose tissue, the exact intermediates between the embryonic stem cell and the mature fat cell, the control of white versus brown preadipocyte commitment, and the control of the relative amounts and functional heterogeneity among white fat cells in different depots. But, adipocytes are thought to be derived in a sequential pathway of differentiation from mesenchymal stem cells (MSCs), which themselves are believed to arise from mesoderm (the middle layer of the three primary germ cell layers in early embryonic stage – between ectoderm and endoderm). However, precise lineage tracing studies have not been performed due to missing definite features of different stages of development on the one hand, and technical issues due to problematic handling of these cell types in standard methods (like FACS, see below) and missing reagents (antibodies against specific surface markers) on the other hand. By definition, MSCs are multipotent stem cells and thus endowed with self-renewal properties and potential to differentiate into all mesenchymal cell types such as adipocytes, osteoblasts, chondrocytes, myoblasts and cells forming connective tissue. It is believed that MSCs give rise to a common precursor (*adipoblast*) which develops into committed white and brown preadipocytes that in turn differentiate into mature adipocytes of different types upon appropriate stimulatory conditions. However, in absence of any known unique gene expression or surface marker or morphological characteristic, it is not clear if separate adipoblasts and/or preadipocytes for white and brown fat cells exist or if there are different white preadipocytes for different WAT depots (Gesta et al. 2007), (Billon et al. 2008). Indeed, there are publications questioning the sole development of mature adipocytes from adipoblasts or preadipocytes. It has been shown that BAT-like tissue can also be derived from myoblastic precursors and fibroblastic cells (skin fibroblasts from mouse and human) by the zinc finger protein *PR domain containing 16* – a bi-directional developmental switch – *in vivo* (the PRDM16-C/EBP-β complex stimulates brown adipogenesis by binding to PPARγ and activating its transcriptional function) (Seale et al. 2008, Kajimura et al. 2009). Also for the white lineage there are very interesting experimental results (from mouse studies) demonstrating that white adipocytes are not necessarily generated from resident adipose tissue mesenchymal progenitor cells but can also arise from the myeloid lineage thus originate from hematopoietic stem cells (HSCs) (The myeloid lineage is the creator of all blood cells except *T-cells*, *B-cells* and *NK-cells* which belong to the lymphoid lineage; but the promiscuity of myeloid cells is displayed by their ability to also generate phenotypes of skeletal muscle (Camargo et al. 2003), vascular endothelium (Bailey et al. 2006), and liver (Willenbring et al. 2004)) (Majka et al. 2010). Majka et al. point out the possibility of a hematopoietic origin for the *de novo* development of a subset of white adipocytes and also for a previously uncharacterized adipose tissue resident mesenchymal progenitor population. These *bone marrow progenitor* (BMP)-*derived adipocytes* and adipocyte progenitors show differences in gene activity as compared with established white and brown cell lines. In essence, global gene expression analysis suggest that BMP-derived adipocytes are true fat cells but differ from conventional white and brown adipocytes in decreased expression of genes involved in mitochondrial and peroxisomal biogenesis and lipid oxidation (suggesting lower oxidative capacity), and increased expression of inflammatory cytokine genes and chemotactic factors. Moreover, the BMP-derived adipocytes accumulate with age, and occur in higher numbers in visceral than in subcutaneous fat. Therefore these cells might account (at least in part) for adipose depot heterogeneity and the negative effects on metabolism and inflammation seen with (especially visceral) obesity and increasing age. Although these findings have been made in a mouse model, the results immediately suggest themselves as possible explanation for the abovementioned observations in humans. The generation of adipocytes from BM myeloid cells is also

interesting with respect to the ability of macrophages (and other myeloid cells) to express certain adipocyte markers including FABP4 (see below) (Pelton et al. 1999), PPARγ (Marx et al. 1998), and hormone-sensitive lipase (Yeaman 2004) (but not *adiponectin* (Luo, Jian Liu, B Hong Chung et al. 2010)). Moreover, macrophages respond to *thiazolidinediones* (also known as *glitazones*), inducers of adipogenesis, by up-regulating adipocyte factors (Pelton et al. 1999). On the other hand, the elevated levels of inflammatory cytokines released by BMP-derived adipocytes may reflect their origin from cells involved in immune function and inflammation. As mentioned, aging is related to fat redistribution from subcutaneous to visceral depots in humans. This pattern of redistribution together with the reasonable accumulation of BMP-derived adipocytes in abdominal depots also in humans, suggests an increasing impact of these cells on visceral fat function with effects on metabolism and physiology, and thereby potentially promoting the development of age-related diseases. Moreover, BMP-derived adipocytes have, compared to other adipocytes, a decreased expression of *leptin*, as mentioned, involved in the regulation of energy expenditure and satiety. Thus, increasing relative amounts may lower *leptin* levels, leading to increased food intake and decreased energy expenditure, explaining the general trend to increasing fat mass with age. There is also a link to *insulin* regulation as the elevated expression of inflammatory cytokines and chemotactic factors may impair *insulin* sensitivity. *Interleukin-6* (IL-6), for instance, is able to directly decrease *insulin* sensitivity via inhibition of *insulin* receptor signaling (Senn et al. 2002) – noteworthy, 30% of circulating IL-6 is released from adipose tissue and this amount increases proportionally with increasing body mass (Gnacińska et al. 2009) . The latter considerations are supported by the findings that in humans visceral adipose tissue has a lower *leptin* production (Woods et al. 2003), higher number of macrophages, and greater inflammatory cytokine production than subcutaneous fat (Hamdy et al. 2006). The relation between macrophages and adipocytes or adipokines, respectively, is also of particular importance as *adiponectin* has been shown to suppress the migration of *macrophages* (or *monocytes*, the precursors) to vascular lesions and their transformation into *macrophage foam cells* in arterial walls. Thus *adiponectin* may have an anti-inflammatory and antiatherogenic role in *atherosclerosis*, what is considered a chronic inflammatory disease and disorder of lipid metabolism – and the leading cause of death in industrialized nations (Lusis 2000, Lau et al. 2005). The relevant morbid process is initiated by the accumulation of cholesterol-rich lipoproteins in the arterial wall resulting in the recruitment of circulating monocytes, their local adhesion, and subsequent differentiation into tissue macrophages. The macrophages for their part accumulate large amounts of lipid to form the foam cells that participate actively in the development of the atherosclerotic lesion (Luo, Liu, Chung et al. 2010).

Model systems

Although the exact origins may still be unclear, the establishment of immortal preadipocyte cell lines that were selected from adipose tissue for their ability to accumulate triglycerides in cytoplasmic lipid droplets has provided several useful model systems for studies on adipogenesis, such as 3T3-L1 cells (mouse) and hMADS (human Multipotent Adipose-Derived Stem) cells. hMADS cells can be directed to differentiate into both white and brown (or at least white-like and brown-like) adipocytes (Rodriguez et al. 2004, Rodriguez et al. 2005, Elabd et al. 2009). They are the primarily used model system in our lab to unveil the regulatory networks driving adipogenesis (both white and brown), with a focus on the role of microRNAs. hMADS cells also provided the basis for the *in vitro* based microRNA-target prediction analysis (see **1.6**).

Regulators of adipogenesis – why the focus on PPARγ makes sense

The transition from preadipocytes to adipocytes involves four stages: growth arrest (upon confluence), clonal expansion (post-confluence mitosis; reentry of cell cycle upon appropriate stimulation; likely necessary for chromatin reorganization to facilitate induction of the adipogenic genes), early differentiation, and terminal differentiation. A transcriptional network coordinating expression of hundreds of proteins responsible for establishing the mature adipocyte phenotype has already been unveiled and is discussed in detail in the reviews (Farmer 2006), (Rosen & MacDougald 2006). The core of this network, with the nuclear receptor PPARγ (a member of the PPAR family) and members of the C/EBP family being the major players (transcription factors, see **1.4.4.1**), is depicted in **Fig. I2**.

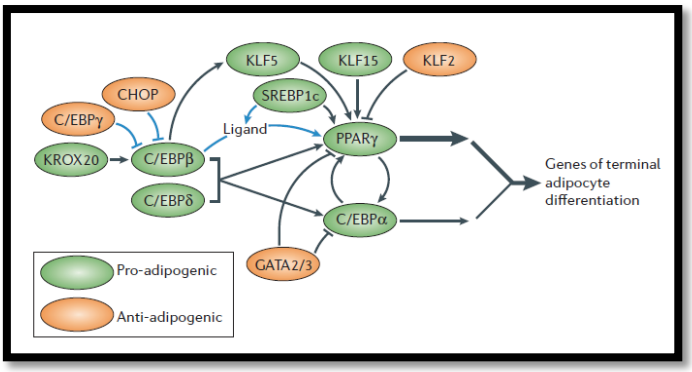


Fig. I2 Core transcription factor network regulating adipocyte differentiation. Black lines indicate effects on gene expression, blue lines on protein activity; adapted from Rosen & MacDougald 2006.

Indeed, *peroxisome proliferator-activated receptor-γ* (PPARγ) is also entitled the 'master regulator of adipogenesis' as it is both necessary and sufficient for adipocyte differentiation. While no other factor has been discovered which would be able to drive differentiation in the absence of PPARγ, forced expression of PPARγ is sufficient to induce it. Consistently, crucial signaling pathways in adipogenesis converge on the regulation of PPARγ expression or activity. For instance, the pro-adipogenic *CCAAT/enhancer-binding proteins* (C/EBPs) and *Krüppel-like factors* (KLFs; a large family of C2H2 zinc-finger proteins that regulate apoptosis, proliferation and differentiation) have all been shown to induce at least one of the two PPARγ promoters, whereas, for instance, the anti-adipogenic GATA factors function in part by repressing PPARγ expression.

PPARγ (also known as *glitazone receptor* or NR1C3 – nuclear receptor subfamily 1, group C, member 3) is expressed as several isoforms (*Ensemble* release 59 states 10 protein coding variants, where some proteins are known to be and some might be structurally identical), generated by alternative splicing and promoter usage of the same gene. The proteins PPARγ1 and PPARγ2 (the ones studied so far and likely the most important forms with respect to general structure similarity and expression levels) also have an identical structure except for additional 30 amino acids at the N terminus of γ2. Both are induced during adipocyte differentiation, but the isoform γ1 is also found in many other cell types (virtually all tissues except muscle), whereas γ2 expression is restricted almost exclusively to adipose tissue and the intestine. In adipose tissue the ratio of PPARγ2 to PPARγ1 has been shown to

increase in obese patients in correlation with their BMI, but the relative roles in adipogenesis remain an open question, as available results are inconclusive. Inhibition of both promoters in mouse 3T3-L1 cells, followed by ectopic expression of either PPARγ1 or PPARγ2 showed that γ2 was able to rescue adipogenesis, whereas γ1 was not. In contrast, ectopic expression of γ1 in $PPAR^{-/-}$ mouse embryonic fibroblasts (MEFs) was sufficient to induce adipogenesis. Furthermore, adipose-selective knockout of PPARγ2 (in mouse) gave rise to *insulin*-insensitive animals with reduced fat (but still substantial amounts of adipose tissue) suggesting that PPARγ1 can compensate for many of the adipogenic functions of PPARγ2 – though, regulating *insulin*-sensitivity (what obviously might be in the job profile of PPARγ2) is a role of major importance (Farmer 2006, Rosen & MacDougald 2006).

In general, nuclear receptors (NRs) regulate the expression of target genes (usually with a need for complementary cofactors) in response to binding of steroid hormones, fatty acids and other ligands (McKenna & O'Malley 2002). Accordingly, unliganded PPARs (referring to the whole family made up of three subtypes: α, β/δ, and γ) heterodimerize with *retinoid X receptor-α* (RXR-α) and represses transcription of the particular target gene when bound to DNA by interacting with corepressor molecules. Upon ligand binding, conformational changes in PPAR cause the dissociation of these corepressors to enable a spatiotemporally regulated association of coactivators and coactivator-associated proteins driving gene expression. Indeed, PPARs interact with very large multisubunit cofactor protein complexes, some exhibiting histone acetyltransferase or methyltransferase activity, while others function as ATP-dependent chromatin remodelers or linkers to the basal transcription machinery. PPARs are known to function as sensors for fatty acids and fatty acid derivatives and to control among others metabolic pathways involved in the maintenance of energy balance (Viswakarma et al. 2010, Yu & Reddy 2007). PPARγ is not only crucial for adipogenesis but also necessary for maintenance of the mature phenotype, where also C/EBPα, the second central transcription factor (**Fig. I2**), may be required, as it seems to maintain expression of PPARγ in mature fat cells. In addition, C/EBPα induces many adipocyte genes directly, and *in vivo* studies indicate its important role in the development of adipose tissue. Although it is not required for accumulation of lipid, it is necessary for the acquisition of *insulin* sensitivity. Indeed, a lack of C/EBPα in terminal differentiation results in *insulin* resistance *in vitro* and an inability to develop WAT *in vivo* (Farmer 2006, Rosen & MacDougald 2006). For a recent review on PPARγ and C/EBPα activities see the work of Siersbaek et. al (Siersbaek et al. 2010) – reporting also the interesting finding that although PPARγ/RXR binding sites are enriched at promoters, the majority of responsive elements are located in regions far away from transcription start sites, a circumstance also known from several other nuclear receptors (based on CHIP data; *chromatin immunoprecipitation*).

Signaling and organelle function – factors that may cause disease states

Excessive energy substrates cause adipose hypertrophy ("fat cell overcrowding") which in turn appears to be associated with abnormalities of adipocyte function, particularly *endoplasmic reticulum* stress, mitochondrial stress and suppression of mitochondrial biogenesis – thereby affecting lipid and *glucose* metabolism, ATP production, and disposal of reactive oxygen species – what may be central to the pathophysiologic effects of obesity and results in consequences such as *insulin* resistance and production of certain adipokines and inflammatory mediators (de Ferranti & Mozaffarian 2008, Bournat & Brown 2010, Rong et al. 2007). Also other publications point out that adipose tissue is not only interlinked with pathways controlling lipid and *glucose* homeostasis but in addition

with more general metabolic and immune response processes. Hotamisligil and some members of his lab (the following citations) show that under conditions of nutritional excess, an implication can be a metabolically driven, low-grade, chronic inflammatory state that targets metabolically critical tissues and thereby adversely effects systemic homeostasis. Moreover, they also report that *endoplasmic reticulum* (ER) stress (meaning dysfunction) is related to metabolic and immune regulation (Hummasti & Hotamisligil 2010, Hotamisligil 2010) – a serious fact since the ER is an organelle with diverse essential functions, among others it is a major protein synthesis and folding compartment for secreted, plasma membrane and organelle proteins. In addition, it is responsible for lipid (especially phospholipid) and membrane synthesis, lipid droplet generation, cholesterol regulation, and Ca^{2+} storage (Calcium is a very important second messenger in many signaling cascades regulating gene expression and enzyme activity.). Obesity additionally seems to be linked with severe downregulation of *autophagy*, a process involved in the degradation of cytoplasmic components including damaged organelles and proteins, thereby also increasing ER stress and impairing *insulin* signaling (Yang et al. 2010). Dysfunction of ER (especially when the negative effects are enhanced by reduced autophagy) results in accumulation of abnormally folded proteins which can interfere with normal cellular functions. A cell typically responds with interventions in regulatory pathways to inhibit protein synthesis and increase clearance of abnormal proteins – this reaction is referred to as *unfolded protein response* (UPR). If ER and cell homeostasis cannot be sufficiently restored, UPR induces apoptosis. In mouse models adipocyte ER stress caused a decreased production of *adiponectin* (per up regulated *C/EBP homologous proteins*; CHOP), providing insight into how intraorganelle/intracellular dysfunction may be communicated systemically via circulating adipokines. The possibility of a relationship between obesity-related ER stress and metabolic dysfunction in obese humans is also supported by the fact that markers of ER stress in adipose tissue significantly decreased and *insulin* sensitivity of hepatic, skeletal muscle, and adipose tissue markedly improved after weight loss (Gregor et al. 2009).

Among the secreted molecules of adipose tissue also so called *lipokines*, such as *C16:1n7-palmitoleate*, do their work in a lipid-mediated endocrine network with effects on *insulin*-sensitivity, for instance (Cao et al. 2008). In general, fatty acids serve not only as energy source but in addition as signals for metabolic regulation and as modulators of gene expression. In this context a special role might be assigned to *fatty acid binding proteins* (FABPs) which are a mostly tissue specific abundantly expressed and evolutionary conserved class of molecules that bind long-chain fatty acids and coordinate lipid responses in cells. The exact biological function and mechanisms of action are still to elucidate but in the broader context they are viewed as lipid chaperones that escort lipids and dictate their biological function. FABPs appear to access the nucleus under certain conditions, and potentially target fatty acids to transcription factors, such as members of the PPAR family. Several FABPs themselves are controlled by these transcription factors, which are known to be liganded by fatty acids or other hydrophobic agonists. *Adipocyte FABP* (A-FABP; FABP4; aP2) acts at the interface of metabolic and inflammatory pathways and has a central role in obesity, *insulin* resistance and atherosclerosis (shown in mouse). In a large population sampling, human individuals with an A-FABP promoter variant causing altered C/EBP binding and thereby diminished A-FABP expression had lower triglyceride levels, a reduced cardiovascular disease risk, and they were protected from obesity-induced type 2 diabetes.

Expression of A-FABP is highly regulated during differentiation of adipocytes by fatty acids, PPARγ agonists and *insulin*. Interestingly, A-FABP is also expressed (approximately 10.000 fold lower) in macrophages, where it coordinates inflammatory activity including production of cytokines such as TNF-α, *interleukin 1β* and *6* (IL1β and IL6), and *monocyte chemoattractant protein 1* (MCP1). In

mouse knockout studies A-FABP$^{-/-}$ macrophages showed a suppression of these inflammatory signaling responses what was highly beneficial against the formation of atherosclerotic lesions. A-FABP is released from adipocytes and abundantly present in human serum, where the concentration may be associated with obesity, diabetes, and cardiovascular diseases. The question on the function of A-FABP in serum and possible effects on distant cells is untreated so far, but a connection with macrophages is conceivable (Furuhashi & Hotamisligil 2008). For some more details and hypotheses on adipose tissue-related signaling see the extensive review published by Lee et. al (Lee et al. 2009).

1.4 Shaping information – the puzzling complexity on small scales

Knowledge of genomic architecture has improved a lot since the first definition of a "*gene*" as *protein-coding unit* was formulated, and it turns out that facts like overlapping transcription units, antisense strand transcripts, alternative transcription start sites, promoter variants, splicing variants (partly spanning extreme distances), transcripts across multiple genes and distal regulatory elements reveal a much higher information content and density than imagined so far, challenging our understanding.

Today some 21.000 protein-coding genes are known in human (*Ensembl* database version 59.37d), whose sequences make up (only) less than 1.5% of the genome. But, data integration and use of evolutionary and computational analysis provide convincing evidence, that the human genome is pervasively transcribed, such that the majority of its bases (some 93% of genomic sequence) can be found in primary transcripts (which are largely unannotated) originating from intergenic, intronic and antisense locations, including the by far largest part (~98%) of so called *non-(protein)-coding transcripts* (*ncRNA*s) often overlapping protein-coding genes on the same (sense) or opposite (antisense) strand of DNA. These ncRNAs include structural RNAs (like rRNAs, tRNAs and snRNAs) and more recently discovered regulatory RNAs (like microRNAs, piRNAs and endogenous siRNAs). Studies of non-coding transcripts of known biological function have begun to unveil a complexity in genome organization prompting a reconsideration of what constitutes the fundamental functional element of the genome and how it relates to phenotypic variation.

Comparing current *Ensembl* data for human (database version 59.37d; known protein-coding genes: 21.257; pseudogenes: 12.599; RNA genes: 8.483; gene exons: 559.206; gene transcripts: 148.792), mouse (database version 59.37I; known protein-coding genes: 22.083; pseudogenes: 5.006; RNA genes: 5.502; gene exons: 387.948; gene transcripts: 87.656) and the worm *C.elegans* (database version 59.210a; known protein-coding genes: 20.224; pseudogenes: 1.552; RNA genes: 16.344; gene exons: 161.896; gene transcripts: 45.625), clearly indicates that the number of protein-coding genes does not (mainly) account for the complexity of an organism – what has been thought for a long time. From genomic analysis it is evident that as an organism's complexity increases, the protein-coding contribution of its genome decreases. A portion of this paradox may be resolved through alternative pre-mRNA splicing (be aware of the exon numbers) and post-translational modifications, both contributing to increased diversity of protein species, but another really important role in the explanation might be filled by ncRNAs and their widespread regulation networks. Of course, also the regulatory networks and signaling cascades constructed by proteins are essential – and amazing. Putting aside the issue of pseudogenes (also not as useless as expected (Khachane & Harrison, Zheng & Gerstein 2007, Suyama et al. 2006)) and the number of RNA genes (which may further increase much more in human and mouse than in *C.elegans* due to different genomic sizes and complexity and hence

scientific progress in these organisms), it is obvious that the number of total gene transcripts is related to the complexity of these organisms, whereby the vast majority of transcripts in human and a high amount in mouse is non-coding and presumably regulatory. It seems pretty much like in architecture, where you can create any building from a plain house to a cathedral using the same kinds of materials but very different plans telling what's to be put where (and when you need it in which amounts).

The mechanisms of action and biological roles of regulatory ncRNAs are extremely diverse (Prasanth & Spector 2007) ranging from mRNA destabilization over modulating translation and transcription right up to epigenetic events, thus they are involved in almost every biological process – including developmental timing, cell differentiation and proliferation, cell death, metabolic control, transposon silencing and antiviral defense. So far, it is not clear how many ncRNA genes are present in the human genome, but one abundant class of ncRNAs which is apparently capable of all these functions has come into the focus of interest about 10 years ago – the microRNAs (Carninci 2010, ENCODE 2007, Prasanth & Spector 2007, Gingeras 2007, Levy et al. 2007, Claverie 2005, Venter et al. 2001, RIKEN et al. 2005).

1.4.1 Introduction to microRNAs

The first representative of this small RNA family, lin-4, was identified in a genetic screen in *C.elegans* in 1981 (Chalfie et al. 1981) and was molecularly characterized in 1993 (Lee et al. 1993), yet the term *microRNA* (or *miRNA*) was introduced as recently as 2001 in a set of three articles in *Science* (Ruvkun 2001).

miRNAs are endogenous single-stranded RNA molecules of 19 to 25 nucleotides in length, which (together with proteins) regulate gene expression on the post-transcriptional level by targeting mRNAs preferentially in the 3'-untranslated regions (UTRs) – and recent data indicates that they are also involved in transcriptional gene regulation. Under certain conditions, some of them are expressed at high levels of up to ten thousands of copies per cell and can thus play important regulatory roles by controlling their messenger RNA (mRNA) targets (Lim et al. 2003). Friedman et al. provide bioinformatic data showing that more than 60% of human protein-coding genes (mRNAs) have been under selective pressure to maintain pairing to miRNAs. In total, within human 3' UTRs more than 45.000 miRNA target sites have been identified as conserved above background conservation levels. This proportion has been determined based on 695 known human miRNAs (miRBase release 10) and might rise with further discovered miRNAs (Friedman et al. 2009).

The recent development of deep sequencing technologies and computational prediction methods has enabled and accelerated the discovery of less abundant miRNAs (and small RNAs in general) (Pantano et al. 2010, Morin et al. 2008) – the workflow for miRNA detection and so far used algorithms are reviewed in (Li et al. 2010) and (Mendes et al. 2009). Briefly, based on characteristics of known miRNAs, filter-based methods are looking for sequences wherein the potential mature miRNAs are included in predicted minimum thermodynamic free energy fold-back precursor structures with extensive base pairing in the miRNA region and as possible void of any large internal loops or bulges. Moreover, many algorithms filter for fold-back structures being phylogenetically conserved and some also make use of the conservation of proximal regions. The drawback of algorithms working on this base is that they can only resemble previously identified miRNAs. Moreover, the high fraction of phylogenetically

well-conserved miRNAs may reflect the bias of search procedures used so far. Indeed these procedures are critically dependent on conservation criteria to attain reasonable levels of specificity.

Later approaches make use of conventional *machine learning methods* and thereby try to generalize from a positive set of previously known miRNAs and negative sets of stem-loops presumed not to be miRNA precursors (the construction of reliable negative sets is obviously a weak point here). Most approaches of that sort are realized by SVMs (support vector machines), HMMs (hidden Markov models), neural networks (NN) or Naïve Bayes (NB) algorithms making use of features concerning sequence composition, topological properties of the stem-loop (e.g. the position of the mature miRNA), and thermodynamic stability. However, since machine learning approaches usually do not incorporate information regarding transcription potential or genomic context, but concentrate on stem-loop features, they may be misclassifying an important portion of the candidates. Furthermore, the positive examples are obtained from miRNAs previously identified by experimental procedures or other computational methods and therefore strongly biased towards highly expressed and/or conserved miRNAs. Nevertheless, with a growing number of miRNAs being identified the performance of machine learning methods can be expected to increase.

As mentioned, latest progresses in deep sequencing techniques were raising the ability to detect and sequence low-abundance transcripts in an unbiased, quantitative way. Computational methods based on this new experimental data have achieved great success in discovering novel miRNAs in the recent past. But, as besides miRNAs also other small RNAs are amplified, more sophisticated algorithms were demanded to sieve out the miRNA transcripts – algorithms like *miRDeep* (Friedländer et al. 2008), for instance. This method is based on a probabilistic model to assess the compatibility of the pattern of sequenced transcripts with the expected one related to properties of miRNA biogenesis – that is, true miRNA precursors should have a characteristic signature with frequent reads of the mature region of the stem-loop and less reads of other parts of the structure (Mendes et al. 2009). In addition, several computational pipelines for processing high-throughput data have been developed (Addo-Quaye et al. 2009, Hackenberg et al. 2009). Zhu et al. provide a listing of so far developed bioinformatic tools, amongst others containing *miRDeep* which is also part of *miRTools* (E. Zhu et al. 2010), a new web application integrating multiple computational approaches and databases.

Despite the usual problems with computational predictions, the number of known miRNAs is continuously increasing. *miRBase* – a repository of miRNAs and miRNA genes from many organisms – states over 1000 miRNAs in humans at present and in total over 15.000 miRNAs in 142 species are known so far ((Griffiths-Jones et al. 2008); *miRBase* release 16). miRNAs can also be found in plants, with on average 120 miRNA-encoding genes (Jones-Rhoades et al. 2006) and in invertebrate animals with about 150 genes (Lai et al. 2003, Ruby et al. 2006). Those identified in one species are often conserved in closely related ones (some of them even from worm to human (M Lagos-Quintana et al. 2001)) but there are also species-specific miRNAs (including human miRNAs not conserved in chimpanzees, for instance (Berezikov et al. 2006)) and there's no conservation between the miRNAs in plants and animals. Generally, it is reasonable that mutations in highly conserved regions would lead to non-viable forms, or forms that are eliminated through natural selection processes. Thus, conservation is seen as hint for the importance of a gene as it is plausible that a conserved sequence has been maintained by evolution despite speciation. Indeed, conservation and characteristic conservation profiles alone may be successfully used in comparative analyses to predict miRNA genes in approaches like *phylogenetic shadowing* (Berezikov et al. 2005) and *target-centered* strategies (Xie et al. 2005).

Many miRNAs have intriguing expression patterns, e.g. stage-specific expression in embryonic development (Caygill & Johnston 2008, Giraldez et al. 2005, Wienholds et al. 2005, He & Hannon 2004, Pasquinelli et al. 2000, Lau et al. 2001), primarily or exclusive expression in certain organs or cell lines (often necessary to establish and maintain the particular cell identity) (Makeyev et al. 2007, Wienholds & Plasterk 2005, Lim et al. 2005, Chen et al. 2004, Krichevsky et al. 2003, Lagos-Quintana et al. 2002) and aberrant expression linked with cancer, diverse diseases, stress and aging (O'Connell et al. 2010, Garzon et al. 2010, Wiesen & Tomasi 2009, Hamrick et al. 2010, Chen et al. 2010, Park et al. 2009, Lee & Dutta 2009, Zhang 2008, Zhang et al. 2007). They also play critical roles in cell cycle control, apoptosis, cell growth and differentiation (Park et al. 2009, Hwang & Mendell 2006, Xu et al. 2003).

What makes this molecule class very interesting beyond the wish to understand cellular function is the fact that as the expression of miRNAs depends on cell type, developmental stage and health or disease, they are likely to work as potent biomarkers and therapeutic targets or agents – even more so since the discovery of their action on the transcriptional level, which implies epigenetic modifications (Suzuki & Kelleher 2009, Kim et al. 2008).

1.4.2 microRNA biogenesis

Unsurprisingly the information needed to build miRNAs resides in the nuclear DNA, but still little is known about transcriptional control of these genes. Biogenesis, structure and function of miRNAs in plants differ in some points from those in animals (Bartel 2004). The scope of this work will be on the conditions in animals and humans, respectively. A deeper interest in the general workflow of the transcription process in eukaryotic cells also has to be redirected to other sources – there're plenty of good text books taking care of that topic (e.g. Stryer L. 2003; see references). Though, it might be advisable to also refer to up to date review papers.

The representation of miRNAs after transcription is called *primary microRNA* (*pri-miRNA*) (Lee et al. 2002). pri-miRNAs have conventional exon-intron composition (but with few introns) and contain local stem-loop structures. A significant fraction of pri-miRNAs have lengths between 1 and 10 kb (the rest beyond) and are therefore much longer than the processed stem loops (**Fig. I3**) currently used to define miRNA genes. The disparity between the length of the transcribed sequence and the final functional product after maturation could indicate an additional function for (at least some) pri-miRNAs (Saini et al. 2008, Ohler et al. 2004). So far, efforts have been focused on the regulatory function of miRNAs, therefore little is known about how the miRNA genes themselves are regulated. The current view is that approximately 50% of mammalian miRNAs are so called *intragenic*, residing in introns and very rarely in exons (mostly in 3' UTRs) of protein-coding genes (with a broad spectrum of molecular functions) and in introns and exons of non-coding transcription units (ncRNAs), though alternative splicing actions determine whether a miRNA is intronic or exonic. All exonic and the majority (~ 80 %) of intronic miRNAs are located in the sense orientation and can thus be transcribed with the host gene. The remaining intronic miRNAs located in antisense orientation have to be part of transcription units with an opposite orientation to the presumptive host gene (Rodriguez et al. 2004, Weber 2005, Saini et al. 2008, Kim et al. 2009).

miRNAs embedded in host genes were long thought to be mainly controlled by the same regulatory elements and thus transcription factors and share the same primary transcript with their host (Baskerville & Bartel 2005a, Bartel 2004). But, Corcoran et al. showed that miRNA genes can be regulated by

promoters located several kb away – a circumstance already known from mRNA regulation – and that these promoters exhibit the same general features like those of protein-coding genes. In addition the group found evidence that a substantial amount (26 %) of the *intragenic* miRNAs may be regulated not by the host genes promoter but by their own (Corcoran et al. 2009). These findings are supported by Monteys et al., who also predicted that ~35% of currently known intronic miRNAs (235 from miRBase 12.0) have upstream regulatory elements consistent with promoter function, with 30% having associated Pol II and 5% Pol III regulatory elements. In one studied case (miR-128-2) even a dual regulation by both intron-resident (Pol III) and host gene (Pol II) promoters is suggested by ChIP data (Monteys et al. 2010). Also Ozsolak et al. found by a combination of nucleosome mapping and chromatin signatures for promoters that one-third of intronic miRNAs have transcription initiation regions independent from their host promoters – the data is supported by poor expression correlations between intronic miRNAs and their host's mRNA (Ozsolak et al. 2008).

Rather little is known about the structure of pri-miRNAs that lie between annotated genes, referred to as *intergenic miRNAs*. They might be part of independent transcription units and thus transcribed by their own promoters, though, some miRNAs that were classified as located outside of known genes could in fact reside in still uncharacterized splicing variants. Most intergenic human pri-miRNA genes measure some 3-4 kb, with distinct 5' *transcription start sites* (TSSs) and CpG islands located within the upstream 2 kb region demarcating the 5' ends, and also poly(A) sites within the downstream 2 kb region of the embedded pre-miRNA defining the 3' boundary (Saini et al. 2007, Saini et al. 2008). Pairwise distance analyses show that 42-48 % of human miRNAs (205 in total, from miRNA registry release 4.0) originate from polycistronic units containing 2-7 pri-miRNAs (42% with a rather stringent maximal distance of 3000 nt, 48% with a distance of at most 10.000 nt), sharing a common promoter (Altuvia et al. 2005). The mature miRNAs coded by these polycistronic transcripts may be either functionally related or distinct (Kim & Nam 2006). Another study (Baskerville & Bartel 2005b) demonstrates that proximal pairs of miRNAs tend to be coexpressed, and that this correlation in expression drops when the distance exceeds 50kb. Also, *expressed sequence tag* (EST) evidence indicates that distant miRNAs may reside on the same transcript. Thus, by adding transcriptional features (TSS, CpG islands and poly(A) signals) to the definition of clusters, the number of clusters may further increase. By this means Saini et al. also predict that around 50% of human miRNAs are polycistronically transcribed in primary transcripts up to 10s of kb long. Interestingly, the length distributions of the two groups of polycistronic and singleton pri-miRNAs are very similar – half of both groups measure between 1 and 10 kb, most of the remainder are 10s of kb long, and a few putative miRNA genes span several 100 kb (though, this may be an annotation bias), whereas the mean length is ~ 4 kb. For comparison, the mean length of protein-coding pre-mRNAs is around 50 kb, and, surprisingly, the mean length of pre-mRNAs containing intronic miRNAs is 150 kb (Saini et al. 2008).

miRNA promoter studies are still rare (Zhou et al. 2007, Ozsolak et al. 2008, Marson et al. 2008, Corcoran et al. 2009). The two RNA polymerases (RNAPol) that qualify for miRNA transcription are RNAPol II and RNAPol III, whereby it is known that the same promoter elements can be used by both polymerases in humans. The promoter region of a class II gene (a gene transcribed by RNAPol II; in this context the most important one, see below) is usually separated into three compartments: (1) a ~100 bp long *core promoter domain* with embedded TSSs for RNAPol binding and binding sites for *general transcription factors* (*GTP*), (2) immediately upstream to this region, a proximal domain several hundred bp long, containing primary regulatory elements and (3) distal domains, thousand of bp

long, constituting secondary regulatory elements (Wang et al. 2009, Juven-Gershon et al. 2008). In general, initiation of transcription is activated by RNAPol binding to its specific target binding sequence within the core promoter. Similar to known protein-coding genes, about 64% of promoters of miRNAs possess CpG islands within the 500 bp proximal region. Also TATA elements are found in 19%, BRE (TFIIB recognition element) in 21%, Inr (Initiator) elements in 47% and DPE (downstream core promoter element) in 87%. Thus, miRNA TSSs identified so far share properties with coding gene promoters and might underlie the same combinatory regulation by *transcription factors* and co-factors (Ozsolak et al. 2008).

RNAPol II is known for producing the mRNAs and several non-coding RNAs (snRNAs and snoRNAs), whereas RNAPol III generates some of the shorter non-coding RNAs, including tRNAs and rRNAs. miRNAs processed from introns of protein-coding genes are mainly transcribed by RNAPol II (see below) and there's evidence that most of the intergenic miRNAs are also RNAPol II products, like the length of pri-miRNAs of more than 1 kb, which is longer than typical RNAPol III transcripts (Lee et al. 2002, Saini et al. 2008, Lee et al. 2004) and the fact that pri-miRNAs often contain repeated runs of uridine residues, what is expected to prematurely terminate RNAPol III transcription (Ohler et al. 2004). The evidence implicating RNAPol II as polymerase of miRNA transcription also includes the discovery that most pri-miRNA transcripts are capped and polyadenylated – unique signature processing characteristics of Pol II transcription products. In addition, pri-miRNA expression levels are greatly reduced by *α-amanitin* at concentrations that specifically inhibit RNAPol II (Lee et al. 2004, Cai et al. 2004, Saini et al. 2008). Due to structural similarity to mRNAs there is also some overlap, where transcripts can function as both pri-miRNAs and mRNAs (Cai et al. 2004). Although these observations suggest that many miRNAs are Pol II transcripts there apparently exist some that are produced by RNAPol III, especially near upstream tRNA sequences and in chromosomal regions with Alu repeats or mammalian-wide interspersed repeats (MWIR) (Borchert et al. 2006, Monteys et al. 2010; Bartel 2004, Turner & Slack 2009). Very recently an interesting form of atypical transcription of some miRNA species has been observed. Gao et al. report findings of transcription upon cellular introduction of DNA fragments (*amplicons*) containing the precursors for known human miRNAs but being devoid of any known promoter sequences. Notably, this transcription appears resistant to conventional inhibitors of RNAPol II and RNAPol III activity and moreover, the different amplicons that were employed do not show any apparent shared sequence features. Based on their experimental data the group concludes that some miRNA precursors may have the ability to self-transcribe their own sequence (Gao et al. 2010).

After transcription of the pri-miRNA, the first step of miRNA maturation is executed by the endoribonuclease *Drosha* RNase III and its cofactor *DGCR8* (*DiGregory syndrome critical region gene 8*), which is an essential supporter for the recognition of the pri-miRNA tertiary structure, acting as determinant for substrate specificity – together they are referred to as *microprocessor complex* (~650 kDa in humans [Gregory 2004]) (**Fig. I4**). Drosha excises a ~60–70 nt imperfect stem loop intermediate with a ~2 nt 3' overhang (which is recognized by the nuclear export factor exportin 5) from the pri-miRNA, known as *precursor miRNA* (*pre-miRNA*) (**Fig.I3**, **Fig.I4**, **Fig.5**).

Fig.I3 *pri-miRNA* stem-loop structure containing the *pre-miRNA* (composed of the colored mature miR-34, the opposing star sequence and a loop)

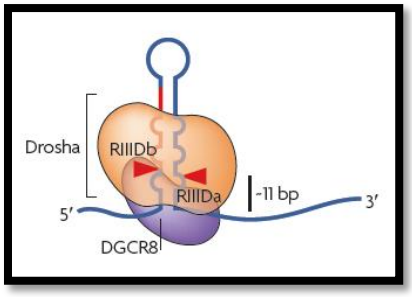


Fig.I4 The Drosha-DGCR8 complex at work adapted from (V Narry Kim et al. 2009)

Although Drosha cuts at specific sites, it does so not by sequence recognition, since there is no common sequence element found in pri-miRNAs. The Drosha-DGCR8 complex recognizes the common structural features among diverse miRNAs (**Fig.I4**). A typical metazoan pri-miRNA stem-loop structure consists of a stem of ~33 bp, with a terminal loop and flanking single-stranded RNA (ssRNA) segments, where the loop is unessential but the flanking ssRNA segments are critical for processing, because the cleavage site is determined mainly by the distance (~ 11 bp, **Fig.I4**) from the stem-ssRNA junction. DGCR8 apparently has the function of a molecular anchor that recognizes this junction and measures the distance (Han et al. 2006, Zeng & Cullen 2005). For more details (like special domains, sites and interactions) on Drosha, DGCR8 and the upcoming Dicer have a look at the nice review by Kim et al. (Kim et al. 2009).

An interesting indication is that apparently intronic miRNAs (embedded in introns of coding and non-coding transcription units) can be processed by Drosha from the *unspliced* transcripts without significantly affecting the production of mature mRNA – suggesting that a continuous intron may not be required for splicing and that the exons are tethered to each other (Kim & V Kim 2007). This finding is supported by Morlando et al. who also show that co-transcriptional cleavage of intronic and intergenic miRNAs by Drosha occurs (Morlando et al. 2008). For some intronic miRNAs (in that context called *mirtrons*), an alternative biogenesis mechanism has been identified. The ends of their stem loop precursor structures coincide with splice sites, they are cleaved by nuclear pre-mRNA splicing through the spliceosome (sometimes followed by additional exonucleolytic trimming on either end) bypassing the Drosha/DGCR8 processing. Besides this exception it seems not unlikely that other non-canonical pathways may be uncovered in future (Berezikov et al. 2007, Ruby et al. 2007, Okamura et al. 2007, Okamura et al. 2008, Kim et al. 2009).

In the next step the pre-miRNA is actively transported from the nucleus to the cytoplasm by the ras-related nuclear protein GTPase (*Ran-GTPase*) and an export protein complex containing a dsRNA-binding export receptor like *Exportin-5* (*EXP5*; in mammals) (Kim 2004, Lund et al. 2004, Yi et al. 2003, YI et al. 2005). The further processing in the cytoplasm is done by another RNase III endonuclease called *Dicer* (**Fig. I8;** also see below, *RISC-loading complex*) (Lee et al. 2003). Dicer was first recognized for its function as generator of *small interfering RNAs* (*siRNAs*) that enable the *RNA interference pathway*

(*RNAi*) (Bernstein et al. 2001) and was later shown to also play a role in miRNA processing (Grishok et al. 2001, Hutvágner et al. 2001, Ketting et al. 2001). Dicer transforms the pre-miRNA form (the stem loop) into the mature miRNA, which is composed of two opposing fragments (a duplex). Besides the positioning, perhaps with particular affinity for a 5' phosphate and the 3' overhang at the base of the stem loop generated by Drosha, Dicer cleaves off the terminal base pairs and the loop of the pre-miRNA irrespective of its sequence (Bartel 2004). So, according to the current model, the specificity and the determination of what's to be both ends of the mature miRNA is the business of Drosha-DGCR8 alone – Dicer is rather promiscuous, it can act on any dsRNA with simple preference toward the terminus and cleavage ~22 nt away (Lee et al. 2003, Bernstein et al. 2001, Elbashir et al. 2001, Zhang et al. 2002, Han et al. 2004).

All RNA-silencing processes – either driven by miRNAs, siRNAs or piRNAs (Piwi-interacting RNAs) – are carried out by large ribonucleoprotein assemblies generally termed *RNA-induced silencing complexes* (*RISCs*). The functional core of every RISC contains a member of the *Argonaute* (*Ago*) protein family. Ago proteins bind small RNAs (19-30 nt), which are used to guide the complex to cognate mRNAs (and non-coding RNAs) through base-pairing interactions. The human genome encodes four Ago (*Ago1-Ago4*; also known as *EIF2C1-4*; the *Ago subfamily*) and four Ago-related proteins united in the so called *Piwi* subfamily (HIWI1-HIWI3 and HILI). Furthermore, several Ago-associated proteins (today more than 15 in humans) including helicases, nucleases and RNA-binding proteins have been identified, raising the possibility of a high degree of combinatorial complexity in RISC composition and thus different modes of function – like translational repression and RNA cleavage, for instance.

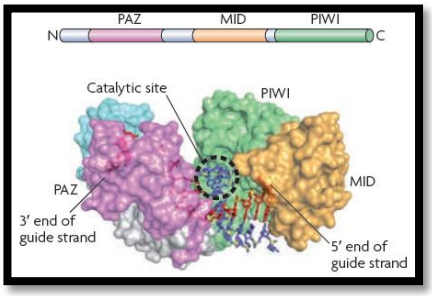


Fig. I5 *Argonaute* protein structure. Ago-family proteins have three characteristic domains: the PAZ domain serves as docking site for the 3' end of small RNAs, the MID domain anchors the 5' terminal nucleotide and the PIWI domain has a structure similar to RNase H (in Ago2 it provides endonuclease activity); translational repression might occur through interactions with other protein factors. adapted from (V Narry Kim et al. 2009)

In humans, at least Ago2 (but probably all family members) is loaded with miRNAs by the action of the so called *RISC-loading complex* (*RLC*), an assembly of proteins comprising Ago2, Dicer and the dsRBD containing proteins *TRBP* (*HIV-1 transactivation-response element RNA-binding protein*) and PACT (*protein activator of interferon-induced protein kinase*). The process of *RISC loading* involves several steps, including pre-miRNA recognition, dicing, guide strand selection, Ago protein loading and second strand removal. In the guide RNA selection usually (but not always) the strand of the RNA-duplex whose 5' end is thermodynamically less stably paired is favored for incorporation into the *Ago* protein complex and is referred to as *miRNA* (*mature miRNA*; *guide strand*), whereas the other strand is referred to as *miRNA** (*miRNA star; passenger strand*). The second strand of the miRNA/miRNA* duplex becomes degraded in most cases (probably by an RNA helicase's activity mediating unwinding and removal of the unselected strand und subsequent exonucleolytic decay; in case of Ago2 by endonuclease activity), but may also be functional sometimes (Schwarz et al. 2003, Sasaki

et al. 2003, Lee et al. 2006, Peters & Meister 2007, Du & Zamore 2007, MacRae et al. 2008). Apparently in cells from certain tissues both strands of a miRNA duplex are utilized, that is, one pre-miRNA can produce two functional miRNAs which target two different sets of genes (Ro et al. 2007).

Once the Ago protein is loaded with a miRNA, it tends to dissociate from the rest of the RLC, allowing for complex formation with other proteins. Van den Berg et al. published a very nice detailed review paper on RISC components, formation and function (van den Berg et al. 2008).

1.4.3 microRNA function

Mature miRNAs' work order is to guide some proteins to the correct target mRNAs (**Fig. I8**), where as mentioned especially the members of the *Argonaute* family and the P-body protein *GW182* (Behm-Ansmant et al. 2006, Eulalio, Tritschler et al. 2009) play important roles. This miRNA-protein complex is often referred to as *miRNA RNP* (ribonucleoprotein particle) or *miRNP*, but also the terms *miRISC* or just *RISC,* originally known from the siRNA-driven RNAi pathway, are used.

miRNPs mediate diverse functions depending on the particular Ago protein and the degree of sequence complementary between the target mRNA and the guiding miRNA (Eulalio, Behm-Ansmant, Schweizer et al. 2007, Eulalio, Behm-Ansmant & Izaurralde 2007, Peters & Meister 2007). Which Ago protein embraces a certain miRNA to form the core of a complex is not clear so far, at least for humans. In flies (*D. melanogaster*) and worms (*C. elegans*) the sorting apparently depends on structural features of the RNA molecule, especially on the degree of complementarity within the RNA duplex (Förstemann et al. 2007, Tomari et al. 2007, Steiner et al. 2007) – this clean-cut mechanism is obviously not at work in humans (Azuma-Mukai et al. 2008).

miRNAs (for simplicity, in fact the true workhorses are the associated proteins as is mostly the case) appear to regulate gene expression in most cases, but not always, through repression, where a single miRNA usually downregulates several targets but in some cases a *direct* downregulation of mRNA levels and/or the production of hundreds of proteins may be caused (Lim et al. 2005, Baek et al. 2008, Selbach et al. 2008) – looking at let-7 as an extreme (because it targets Dicer, as shown a key player in the miRNA pathway), even several thousand protein levels are manipulated (Selbach et al. 2008). This so called *post-transcriptional gene silencing* (*PTGS*) usually takes place in the cytoplasm, though not exclusively as functional RISCs have also been verified in the nuclear compartment (Robb et al. 2005). miRNAs typically hit their target mRNAs (**Fig. I6**) in the 3' *untranslated regions* (*UTR*s) by forming usual Watson-Crick pairs, but there are also predictions and some evidence (based on artificial sites reporter constructs (Kloosterman et al. 2004) and genome wide analysis (Baek et al. 2008, Grimson et al. 2007)) that the *open reading frames* (*ORF*s) and 5' UTRs may contain functional binding sites, though in smaller number in ORFs and much smaller number in 5'UTRs and moreover, those sites seem to be less effective (Bartel 2009). A possible explanation for that is the clearing activity of the translation machinery. Since ribosomes attach to the cap-structure at the 5' end and migrate along the mRNA over to the 3' UTR, it is likely that silencing complexes bound upstream of the 3' UTR are rather displaced by the ribosomes than the ones bound to 3' UTRs. This notion is supported by the observation that the transition to more effective and also more conserved target sites is not at the stop codon but instead occurs around 15 nt into the 3' UTR, as expected if the first few nucleotides were cleared of silencing complexes when the ribosome approaches the stop codon (Grimson et al. 2007, Bartel 2009).

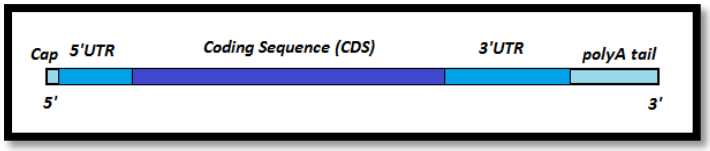


Fig. I6 Structure of a typical eukaryote mRNA. For description see main text.

Eukaryotic messenger RNAs have a typical architecture (**Fig. I6**). The 5' Cap structure is necessary to protect the 5' end against phosphatases and nucleases and thus increases mRNA stability. Moreover, it functions as nuclear export regulator, and enhancer of translation. Caps consist of a very unusual 5',5' triphosphate linkage between a guanine nucleotide and the mRNA. Subsequent methylation of the guanine on N-7, and in some cases of neighboring ribose hydroxy-groups, completes the formation. The *coding sequence* bears the information eventually being translated into amino acid sequences, and the polyadenylation (*polyA tail*) on the 3' end has basically the same functions as 5' capping, it influences nuclear export, translation and stability of the molecule (whereat the polyA tail is shortened over time, and eventually the mRNA is enzymatically degraded) (Guhaniyogi & Brewer 2001). The most important parts of an mRNA molecule, with respect to miRNA regulation, are the untranslated regions (UTRs). The 5'UTR spans the region between the transcription start site (TSS) and the translation start codon (typically hundreds of nucleotides). It bears binding sites for proteins regulating stability and translation of the mRNA (e.g. ribosome binding site), non-protein-related regulatory elements such as *riboswitches* (responsive for low molecular metabolites), and sequences that influence translation by secondary structure formation (Pickering & A. E. Willis 2005). The 3' UTR starts immediately after the translation stop codon and stretches till the polyadenylation signal (up to several kilobases). Besides the polyA signal it contains binding sites for proteins affecting location in the cell and mRNA stability, where AU-rich elements (AREs; see below) may be of particular importance as with their aid, dependent on the bound ARE-binding protein, the mRNA can be stabilized or destabilized. Another class of sequences is that of microRNA binding sites, clearly the most important in the context of the present work. The length of 3' UTRs increases with evolutionary age and organism complexity (with human mRNAs having the longest 3' UTRs) (Mazumder et al. 2003, Pesole et al. 2001). These longer 3' UTRs may provide more regulatory elements that may contribute to a more complex posttranscriptional regulation of mRNAs in humans (Mazumder et al. 2003, Chatterjee & Pal 2009, Stryer 2003, Campbell 2003).

There is evidence that 6 to 8 nucleotides (6-mer to 8-mer binding) at the 5' end of a miRNA (position 1-8) are essential for target site recognition - this sequence has been designated the *seed region* – and that there may occur perfect and imperfect seed pairing with the target (**Fig. I7**) (Lai et al. 2003, Lewis et al. 2003a, Jackson et al. 2006, Rajewsky 2006a, Gaidatzis et al. 2007, Bartel 2009). The sequence surrounding the 3' UTR target region that is complementary to the miRNA seed region (the seed context) also contributes to the function of a miRNP. Features found to influence site efficacy include a position within the 3' UTR at least 15 nt from the stop codon, positioning away from the center of long UTRs, AU-rich nucleotide composition near the site or other measures of high site accessibility, and proximity to sites for coexpressed miRNAs (within 40 nt, but no closer than 8 nt) (Grimson et al. 2007, Nielsen et al. 2007, Saetrom et al. 2007, Bartel 2009). However, it is likely that other important parameters for functional miRNA-mRNA interactions remain to be discovered.

For instance, Lee et al. recently identified motifs in the 5'-UTRs of mRNAs specific to 3' UTRs of miRNAs. Indeed, many miRNAs with these 3'-end interaction sites in 5'-UTRs simultaneously contain the typical 5'-end interaction sites in the 3' UTRs (this miRNA target class is referred to as *miBridge*). The study clearly shows that combinatory interactions between a single miRNA and both end regions of an mRNA can significantly enhance regulatory efficacy, probably by large steric hindrance against ribosome scanning caused by miRNA-associated proteins – for example, the extent of hsa-miR-34a mediated effect on AXIN2 depends on both UTR sites (Lee et al. 2009).

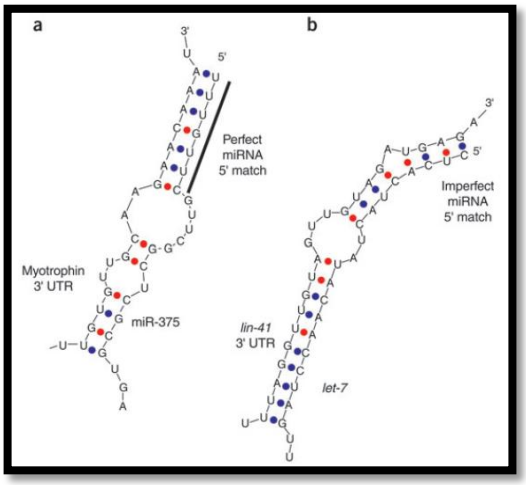


Fig. I7 There exist at least two classes of miRNA target sites in animals. (a) Class 1 targets have perfect, consecutive Watson-Crick base pairings between the 5' end of the miRNA and the 3' UTR of the target RNA and do not require significant further complementarity. (b) Class 2 targets have an imperfect miRNA 5' match but significant complementarity of the remainder of miRNA sequence (referred to as 3'-compensatory sites).
adapted from (Rajewsky 2006b)

Most mammalian miRNAs have multiple isoforms. For instance, the human genome has 12 loci for let-7 family miRNAs. Theses paralogues often have identical seed sequences, thus they are thought to act redundantly. However, because the remainder of miRNAs may also contribute to target binding and because the expression patterns of these sister miRNAs are often different from each other, members of the same seed family may have distinct roles. Hwang et al. demonstrated that despite their small size (at least some) miRNAs contain additional sequence elements (e.g. 3' terminal motifs) that control their post-transcriptional behavior, including their subcellular localization (Hwang et al. 2007, Ventura et al. 2008).

Moreover, the ends of miRNAs are often heterogeneous. Sequence variations are found at both 5' and 3' ends, with a higher variability of 3' ends. The 3' end often contains untemplated nucleotides (nonrandom and evolutionary conserved), which must be added after processing by unknown transferases. Changes in the 5' terminus (maybe due to imprecise or alternative RNase III processing) result in shifts of the seed sequences, which alter the target specificity of the miRNA. Also deletions on both ends are common, probably due to exonucleases. All these variants produced by 5'- and 3'-trimming, nt-substitution (see below, miRNA editing) and 3'-addition are collectively referred to as *isomiRs* (Azuma-Mukai et al. 2008, Morin et al. 2008, Pantano et al. 2010, Kim et al. 2009).

Apparently most miRNPs with a nearly perfect complementarity between guide miRNA and target mRNA mediate mRNA degradation initiated by a cleavage reaction (similar to RNAi mediated by *siRNA/siRISC*) and therefore have to contain the protein Ago2 (the only one in humans/mammals with endonucleolytic ability) – but, in animals this extensive complementarity with consequent cleavage is rather unusual, in contrast to plants (Bartel 2009). RNPs with a greater degree of mismatch

(talking about the entire RNA strand) are more likely to inhibit translation and/or to destabilize and degrade the mRNA (by acceleration of deadenylation and decapping (Eulalio, Huntzinger et al. 2009) and subsequent exonucleolytic mRNA decay) and/or trigger the transport of mRNAs to *mRNA-processing bodies* (also termed *P-bodies* or *cytoplasmic GW-bodies*) (Du & Zamore 2007).

P-bodies contain a variety of proteins including members of the Ago/Piwi family, members of the GW-protein family, RNA helicases and the components of the RNAi and mRNA decay machineries such as decapping proteins (DCP1, DCP2), translational repressors, deadenylase complexes (CCR4:NOT) and several RNA-binding proteins. The presence of P-bodies is considered to be a consequence of RNA-mediated mRNA regulation, but miRNA or more generally small RNA-mediated regulation does not necessarily require localization to P-bodies, actually these bodies are not always detectable (Parker & Sheth 2007, Eulalio, Behm-Ansmant & Izaurralde 2007, Behm-Ansmant et al. 2006). There is also evidence that P-bodies serve as temporary refuge for at least a subgroup of miRNA-repressed, translationally quiescent mRNAs, where miRNA-mediated repression is reversible, with the mRNA being shuttled between P-bodies and actively translating polysomes (Bhattacharyya et al. 2006).

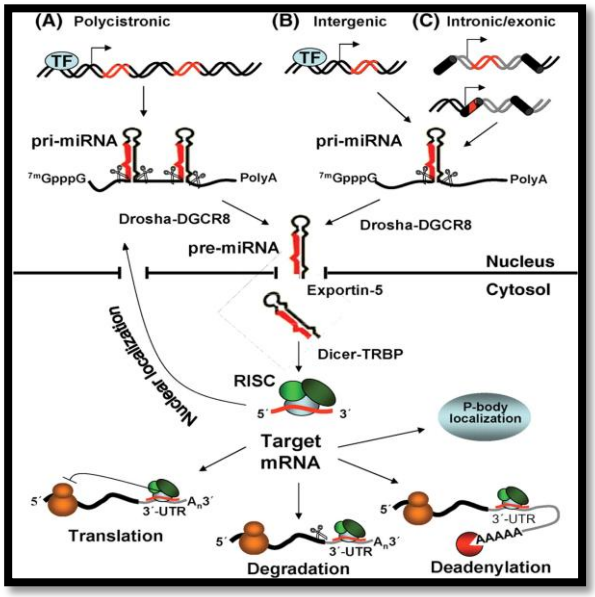


Fig.I8 The current model of (*canonical*) biogenesis and function of miRNAs in terms of post-transcriptional gene silencing (omitting some known additional *regulatory* proteins such as ADAR (Kazuko Nishikura 2010), p53 (Hiroshi I. Suzuki et al. 2009) and others). Sequences encoding miRNAs are shown in red, exons in black and introns in grey. For details see main text. adapted from (Fazi & Nervi 2008)

So far, it was not possible to establish a unified model of miRNA-mediated translational repression, mainly because of (at first sight) contradictory results. Several models have been proposed assuming that miRNAs block initiation at different stages (meaning miRNPs interfere with some proteins that need to bind to the mRNAs cap structure – like eIF4E (Kiriakidou et al. 2007, James et al. 2010) and eIF6 (Chendrimada et al. 2007) or elongation. In addition enhanced decay processes like induction of rapid proteolysis of nascent peptides or deadenylation of target mRNAs followed by decapping and degradation – what can also occur in mentioned P-bodies – are reasonable (Iwasaki & Tomari 2009). It is a fact that different miRISC variants, which can be distinguished by the incorporated Ago proteins, perform different tasks and also employ distinct mechanisms for translational repression (Iwasaki & Tomari 2009) – as is obvious for instance in the case of Ago2 and slicing activity in mammals. But as to that also the bigger picture has to be seen, in which Ago proteins act as cofactors or 'couplers' between the target and the actual craftsman, who may be called GW182 or LIMD1, for instance (Behm-Ansmant et al. 2006, James et al. 2010).

Eulalio et al. provide evidence that miRNAs increase the accessibility of general mRNA degradation enzymes to the mRNA 5' cap structure and the poly(A) tail and trigger mRNA decay by inducing changes in the *mRNP* composition and/or structure, rather than by directly interfering with binding and function of ribosomal subunits. But although deadenylation is a widespread effect of miRNA regulation, it is not absolutely required for silencing, yet when it occurs, the inhibitory effect of miRNAs is strongly enhanced. Whether or not miRNAs trigger degradation is likely to depend on the specific proteins associated with a given target and/or on some specific features of the miRNA-binding site and its RNA context (Eulalio, Huntzinger et al. 2009).

The arising question on what mechanism is predominant, translational repression or mRNA decay, was addressed by Baek et al. and Selbach et al. . Both groups utilized *microarrays* to measure mRNA levels and the *SILAC* (*stable isotope labeling with amino acids in cell culture*) approach to determine corresponding protein levels, after significantly increasing the active amount (by transfection of *miRNA mimics*; see Materials) or decreasing the amount (by transfection of *miRNA antisense inhibitors*, see Materials) of certain miRNAs in the studied cell lines. Minimizing the possibility of false interpretation by utilization of seed-match searches (and thereby more or less ruling out indirect effects on protein levels) one can say that regulation by miRNAs mainly occurs on mRNA level. But there are exceptions to this rule, since a few protein levels changed without any variation in the abundance of their transcripts. In particular, translational repression seems to be more frequent for mRNAs translated at ribosomes associated with the endoplasmic reticulum rather than the cytosolic ribosomes (Baek et al. 2008, Selbach et al. 2008, Esslinger & Förstemann 2009). Moreover, translational repression seems to be reinforced in *miBridge* targets (Lee et al. 2009).

Under certain conditions, like contact inhibition or cell cycle arrest (*quiescence*) due to serum starvation (and likewise occurring during inflammation, angiogenesis and differentiation), miRNA-protein complexes can also activate translation of their target mRNAs, while they normally repress translation of the same target mRNAs in proliferating cells. This up-regulation of translation seems to require *AREs* (*AU-rich elements*) in the 3' UTRs of target mRNAs (what is the case in approx. 12 % of mammalian mRNAs) and miRNPs containing the proteins *Ago2* and *FXR1* (*fragile X mental retardation-related protein 1*) – the latter apparently replacing GW182, an essential partner in the repression complex (Iwasaki & Tomari 2009, Vasudevan & Steitz 2007, Vasudevan et al. 2007, Steitz & Vasudevan 2009).

Based on previous studies with siRNAs, recently other functions besides *post-transcriptional gene silencing* (*PTGS*; the better characterized RNAi pathway) have been related to miRNAs, namely RNA-mediated complex induced *transcriptional gene silencing* (*TGS*; a second distinct RNAi pathway) (Kim et al. 2008, Suzuki & Kelleher 2009) and also *transcriptional activation* (Place et al. 2008) – a form of the so called *RNA activation* (*RNAa*), already observed with dsRNAs (Janowski et al. 2007). Both mechanisms are grounded on sequence complementarity between the miRNA and a target site in the promoter of the regulated gene. As experimental data (Daniel H Kim et al. 2006) led to implication of RNAPol II in small RNA-mediated TGS, two competing models are conceivable, the RNA-DNA model, in which RNAPol II is melting the DNA duplex at the promoter, making one strand available for hybridization with the RNA – enabling a *trans*-regulatory role of miRNAs (and complexed proteins; *RITS*, see below) similar to transcription factors, or the RNA-RNA model, in which RNAPol II synthesizes *antisense or sense transcripts* of the targeted promoters (frequently non-coding transcripts that overlap mRNAs). These transcripts are recognized by miRNAs – or more generally small RNAs (sRNA) – and function as recognition motifs to direct epigenetic silencing complexes termed *RNA-induced transcriptional*

silencing complexes (*RITS*; containing the sRNA, an Ago member and other proteins – in humans yet to elucidate) to the corresponding promoters – pretty much like an extended version of the PTGS mechanics. Recent findings support the RNA-RNA model (Han et al. 2007, Gonzalez et al. 2008, Schwartz et al. 2008).

Kim et al. reported that miRNAs could also act as *cis*-regulators to modify gene expression. At least in the case of the studied miR-320, which is encoded in the antisense orientation within the promoter of the silenced gene, the cis-regulatory role is reasonable. Furthermore, the group provides indications that – similar to translation activation in the PTGS pathway – TGS depends on the cell cycle in that case (Kim et al. 2008).

To date, the components shown to be involved in the small RNA-mediated TGS pathway are the Argonautes Ago1 and Ago2, the DNA methyltransferase DNMT3a and the histone methyltransferase EZH2. RNA-directed TGS seems generally (no matter which type of *ncRNA* facilitates it) associated with epigenetic remodeling events in the form of *DNA methylation* (*RNA-depended DNA methylation; RdDM*) or induction of biochemical modifications of certain residues in the histone tails within the target promoter region (changes of the so called *histone code;* in humans especially RNA-directed methylation of lysine residues 9 and 27 on histone 3 – H3K9me2 and H3K27me3, respectively), which are in turn associated with regional repressive chromatin structures (heterochromatin). But histone modifications can be both activating and deactivating, depending on the type of modification, and the specific residue modified, what matches to the observation that identical RNA duplexes (though certainly complexed with a different set of remodeling proteins) can induce either transcriptional activation (RNAa) or transcriptional silencing (TGS, RNAi) depending on certain circumstances, like basal expression levels (of the regulated mRNAs) which are influenced by growth conditions of the examined cells (Schwartz et al. 2008) or the balance in the bidirectional transcription (of mRNA and related non-coding antisense transcript) (Morris et al. 2008), for instance (Suzuki & Kelleher 2009, Hawkins & Morris 2008).

Although long-term gene silencing is expected in TGS due to epigenetic events, it is interesting and worth mentioning that in *C.elegans* one dose of siRNA was capable of modulating gene silencing that was inherited indefinitely in the absence of the original siRNA trigger (Vastenhouw et al. 2006) – a fact that could be useful in the development of therapies.

1.4.4 microRNAs and transcription factors

1.4.4.1 Transcription factors

Transcription factors (TF) are a class of proteins that bind to specific DNA sequences (the so called *cis-regulatory elements*, *transcription factor-binding sites* or *response elements*) with the objective to control the rate of information transfer from DNA to RNA in temporal and spatial patterns. This control is achieved by stabilizing or blocking the binding of RNA polymerases by electrostatic and geometric interactions and by catalysis of acetylation/deacetylation reactions on histon proteins what weakens/strengthens the association of DNA and histons, making the DNA more/less accessible to transcription. Besides this function the basic defining feature of TFs - that confines them from other proteins, especially other gene regulatory proteins like methylases, deacetylases or chromatin

remodelers for example - is that they contain one or more DNA-binding domains (DBD). TFs perform their regulatory function either alone by promoting (as an *activator*) or blocking (as an *repressor*) the recruitment of RNA polymerases to their binding domains in the promoter region or most often in form of complexes with many other proteins (*coactivators* and *corepressors*) or other transcription factors. As they are essential for the regulation of gene expression they can be found in all living organisms, whereas the number of TFs found in an organism correlates with genome size and larger genomes tend to have more TFs per gene. Many TF binding sites (TFBS) are located in clusters called *cis-regulatory modules* (*CRM*s) that are generally a few hundred base pairs in size, and may interact with more than 20 TFs (Stryer 2003, van Nimwegen 2003, Wang 2005, Venter et al. 2001). TFBS of pri-miRNAs are often found in clusters within the upstream 2 kb region of the pre-miRNAs, and also many TSS and CpG islands lie within 2 kb of the precursor, but a smaller number appears to be 10s of kb upstream (Saini et al. 2007).

Key to the entire transcription initiation, protein recruitment, and formation of activating or repressing protein complexes in the right places are the cis-acting sequence domains embedded either in the proximal core promoter regions or in the primary and secondary distal locations, and functionally manifested as either enhancing or repressing elements of transcription (Sandelin et al. 2007). The common view of transcriptional regulation involves five types of *cis*-regulatory elements – promoters, enhancers and silencers (Arnosti & Kulkarni 2005), insulators (sequences located between enhancers and promoters of adjacent genes, preventing an enhancer from inappropriately activating the promoter of a neighboring gene) (Gerasimova et al. 2000), and locus control regions (LCRs) for long-range interactions (Li et al. 2002), (Spilianakis et al. 2005). However, the current view is known to be overly simplified. For instance, the little information content of TFBS caused by their short length (typically 6 to 10 bases) and their degenerated *consensus sequences*, implies that possible binding sites are present all over the genome numerous times, with the great majority of these not participating in transcriptional regulation. More up to date bioinformatic approaches for TSS and TFBS prediction successfully utilize chromatin structure data rather than using *position weight matrices* (PWMs) and thresholds alone, thereby reducing the false positive rate. It is reasonable that also in the biological system the chromatin structure largely determines whether a certain sequence (motif) has a regulatory role at a certain position, as a 100 to 130 bp window lacking nucleosome binding is usually found to surround active TSSs (Ozsolak et al. 2008, Ozsolak et al. 2007). Another open question is how the different signals from distal regulatory elements can be coupled – what apparently happens by extensive repositioning of genetic loci (containing the *cis*-regulators), allowing dynamic intra- and even interchromosomal interactions (Spilianakis et al. 2005) – without affecting the activity of all neighboring genes (ENCODE 2007).

1.4.4.2 Shared principles

Many TFs bind cooperatively to their specific DNA sequences and/or cooperatively recruit additional transcriptional cofactors (Hobert 2008). Cooperative action is also seen with miRNAs, where multiple target sites in the same 3' UTR can potentially increase the degree of translational suppression *in vitro*. The impact on a target may also be determined by the potential for *several (distinct)* miRNAs to mediate cooperative effects by targeting the same transcript (Doench & Sharp 2004, Krek et al. 2005). However, it is unclear if different miRNAs act *in vivo* in synergistic fashion, it could also be that sites

for different miRNAs in the same 3' UTR merely indicate that the particular mRNA is regulated by different miRNAs in different tissues or at different time points in development (Rajewsky 2006b).

Analyses of microarray data have shown that with most site configurations (meaning seed match types, see (Grimson et al. 2007)), the increased response observed for messages with multiple binding sites is nearly the same as that expected if each site contributes independently to repression. That is, the response of a gene with multiple sites matches that when multiplying the responses for each site working on its own (Grimson et al. 2007, Nielsen et al. 2007, Saetrom et al. 2007). This multiplicative effect is a hallmark of independent action and was also observed with reporter assays (Doench et al. 2003). Although not cooperative in the biochemical sense, the responses to multiple binding sites can (but don't have to) add up substantially. For instance, a message with eight binding sites for coexpressed miRNAs would be repressed by ~25-fold if each site independently decreased the protein output by a third ($0{,}67^8 = 0{,}04$). Exceptions to the overall tendency of independent action have also been found. Two binding sites that are close together (within 40 nt, but no closer than 8 nt) tend to act cooperatively, leading to marked enhancement of repression over that expected from the independent contributions of the two sites (Grimson et al. 2007, Saetrom et al. 2007). By analogy to transcription factors, cooperative miRNA function would provide a mechanism by which their function can become more sensitive to small changes in miRNA expression levels.

Other factors like post-transcriptional modification and editing steps and especially the regulation of binding site accessibility also play important roles in the regulation networks (**Fig. I9**).

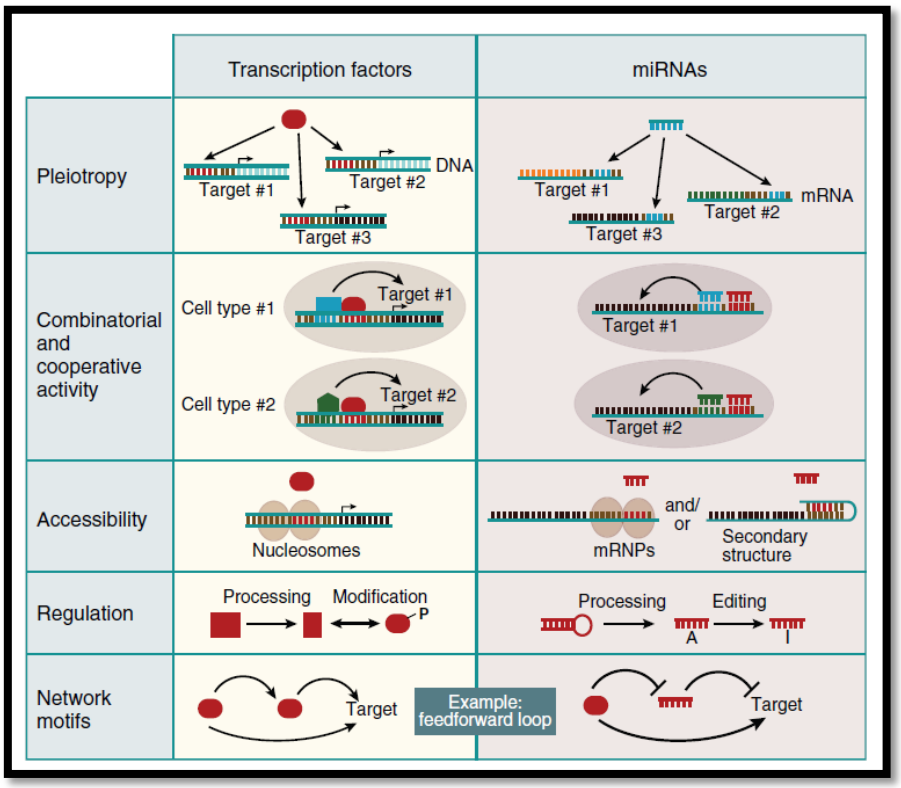


Fig. I9 Visualization of some shared principles of transcription factor and microRNA function.
adapted from (Hobert 2008)

A very important mechanism in post-transcriptional regulation of TFs is carried out by phosphorylation and dephosphorylation reactions, which change their behavior in terms of DNA

binding, activator/repressor activity and protein:protein interactions – or in simple phrase: switching them on or off. Post-transcriptional control of miRNAs reaches from reduced accessibility of pri-miRNAs caused by the NF90-NF45 protein complex (nuclear factor 90 and 45 bind pri-miRNAs and inhibit processing to pre-miRNAs) (Sakamoto et al. 2009) over regulation of the maturation process of miRNAs per influences on Drosha (Thomson et al. 2006) and Dicer (Obernosterer et al. 2006) – enabling the ability of a cell to tune the amount of active mature miRNAs on a layer beyond transcription factors – to RNA editing steps (Kawahara et al. 2007a, Kawahara et al. 2007b, Öhman 2007) known from mRNA processing. Kawahara et al. provide interesting evidence that the edited isoform of a miRNA silences specifically a different set of genes, adding additional complexity to the system – similar to mRNA processing. Moreover, editing of the precursor can interfere with miRNA processing by making them poor substrates of the RNase III proteins.

Besides the controls at transcriptional and post-transcriptional level – regulating miRNA biogenesis, mature miRNA abundance apparently is also a function of actively controlled turnover, with RNA duplexes (of guide/passenger strands) being quite stable and seemingly not subject to degradation. miRNA-Ago complexes – like human siRNA-Ago complexes (Martinez & Tuschl 2004) – are also highly stable, thus, for the termination of miRNA activity proteins (XRN-2 in *C.elegans*, for instance) are necessary to dislodge the miRNA from the complex and subsequently degrade it (in its single-stranded form). This mechanism is modulated by the target mRNA binding status of the miRNA, as mRNAs can stabilize their cognate miRNAs – suggesting a coordination of miRNA and target levels, permitting miRISC reprogramming (loading of a different miRNA and maybe a change in protein composition) when target abundance is low. This additional layer of regulation of miRNA activity has been shown in *C. elegans* but probably it is a general mechanism in animals, which might be important for rapid changes of miRNA expression profiles during developmental transitions and for maintenance of steady-state concentrations of miRNAs (Chatterjee & Grosshans 2009).

Site accessibility is one of the really nice solutions used in biological systems to put things in order. TF binding can be blocked at condition-specific targets by (de)stabilizing nucleosomes at TF binding sites, with nucleosome positioning and remodeling being regulated processes (Buck & Lieb 2006). The interesting question of how the remodeling factors themselves gain access to the nucleosomes or histons respectively was treated by Li and Widom (Li & Widom 2004). Their data suggest a model where dynamic equilibrium conformational transitions in nucleosomes lead to a temporarily increase of the distance between the nucleosomal DNA and the histon core. This spontaneous exposure of DNA target sites allows any protein, both energy-dependent ATP machines and passive binders, to access even buried stretches of nucleosomal DNA. Once recruited to a particular nucleosome, a remodeling factor can move or disassemble the nucleosome, allowing the site-specific binding protein (TF) to bind with much higher equilibrium occupancy than would be possible without these modifications. Schones et al. produced a nice paper on nucleosome positioning, reorganization and modification around *transcription start sites* and *enhancers*, providing also a bunch of references on this topic and the backgrounds (Schones et al. 2008). The accessibility of miRNA targeted recognition sites within mRNAs is also controlled by proteins, RNA-binding proteins (RBPs) such as *HuR* or *Dnd1* often related to AU-rich or U-rich elements, for example. These RBPs counteract miRNP binding to target mRNAs, however, it's unclear how the mechanics behind really works, it may also be, that miRNPs are not blocked by competitive RBPs but influenced in their activity (Kedde & Agami 2008).

One interesting fact was revealed by Cui et al., who found that genes with more TF-binding sites have a higher probability of being targeted by miRNAs and have more miRNA-binding sites on average.

This observation indicates that genes with higher cis-regulation complexity are more coordinately regulated by TF at the transcriptional and by miRNAs at the post-transcriptional level (Cui et al. 2007).

1.4.4.3 Signaling networks – the crosstalk between microRNAs and transcription factors

Regulating regulators, biological systems' joker. Whenever it comes to information processing in biological systems things usually get quite complex due to the formation of networks, which in the majority of cases consist of more players and interactions than we can handle at once.

In the case of TFs and miRNAs we are also talking about networks in which sets of combinatorial expressed TFs and miRNAs delineate individual cell types and cell states by shaping specific gene regulatory programs. The underlying well-defined network motifs, like positive and negative feedback and feedforward motifs (**Fig. I9**) causing system behavior such as signal amplification, dampening and oscillation, are known from control theory. Also interaction patterns matching *Boolean* logic elements like AND- and OR-gates known from mathematical logic and digital electronics can be found – for instance, when looking at the TF binding to promoter regions. As many TFs themselves are subject to miRNA regulation and the expression profiles of miRNAs are very likely basically results from conventional TF-dependent transcriptional control mechanisms (at least when talking about pri-miRNA levels), expectedly miRNAs and TFs are also linked to one another in regulatory networks (Tsang et al. 2007), (Johnston et al. 2005), (Mangan & Alon 2003). Clearly, in the complex data processing network established in a cell, miRNAs not only act on the basal layer of transcription factors but on different levels of signaling cascades. Inui et al. deal with this issue in a detailed review on miRNAs and signaling pathways (Inui et al. 2010).

1.5 microRNA targets – the art of prediction

As mentioned, miRNAs also seem to target promoters or transcripts thereof in the course of TGS. The objective of the present work is the validation of mRNA targets of miRNAs (PTGS), thus when talking about *targets* hereafter one has to think of mRNAs.

In the absence of a clear physical model for miRNA action in animals, most bioinformatic approaches are based on the few principles known from experimentally confirmed miRNA:mRNA pairs. Thus, they seek conserved sites in the 3' UTR with favorable thermodynamic hybridization energies and use the detection of seed matches as primary filter. Other strategies make use of machine learning methods, trying to grasp the general rules of target site recognition from a set of known validated RNA pairs. As experimental identification of miRNA targets is difficult, a wide range of algorithms has been developed. In the following some of them are described briefly, among prominent representatives (e.g. *TargetScanS*, *miRanda* and *PicTar*) also those used in an in-house conducted study with the objective to generate a data set of possible miRNA:mRNA pairs related to obesity. For a more detailed listing and description see the reviews by Li and Mendes (Li et al. 2010, Mendes et al. 2009).

1.5.1 Selected target prediction algorithms

miRanda (Enright et al. 2003)], *MicroInspector* (Rusinov et al. 2005), *DIANA-microT* (Kiriakidou et al. 2004), and *TargetScan* and *TargetScans* (Lewis et al. 2003b) belong to the earlier developed methods. All of them are mainly based on the to date established characteristics of miRNA:target pairing. Of course, they vary in some details, e.g. the exact location of the seed match (usually, position 2-7 or 2-8 of the miRNA), and whether only Watson-Crick base pairs or also G-U base pairs are allowed. But in general all these algorithms check whether the 5' seed of a miRNA is complementary to the 3' UTR of a potential target mRNA and if so, the conservation (at least in closely related species) is examined. Another very important filter (besides conservation) in this type of algorithms is based on thermodynamics. To be thermodynamically preferred, the RNA-RNA duplex formed by the miRNA and its target must have a higher negative folding free energy than the secondary structure of the mRNA.

The more recently developed methods are often based on machine learning approaches, thus being somewhat more sophisticated compared to the first generation, but most of these algorithms also rely on seed matching (or at least partial complementarity) and thermodynamics – however, in many cases conservation filtering (which certainly introduces a bias) could be dropped.

A popular algorithm is *PicTar* (Krek et al. 2005). *PicTar* is scanning aligned 3' UTRs for seed matches to miRNAs followed by a filtering according to their thermodynamic stability. Each predicted target is then scored by using a HMM (hidden Markov model) maximum-likelihood fit approach. Thereby *PicTar* is capable of accounting for synergistic effects of multiple binding sites of one miRNA or several miRNAs acting together. Additionally it provides appropriate scoring of overlapping sites and background-binding.

As experimental studies (Long et al. 2007) suggest that site accessibility is a very critical factor for miRNA binding, algorithms like *MicroTar* (Thadani & Tammi 2006) and *PITA* (probability of interaction by target accessibility) (Kertesz et al. 2007) have been designed to exploit this finding – both are able to predict conserved as well as non-conserved targets and both showed measureable improvements over other prediction methods.

Rna22 is a pattern-based approach (patterns are derived from a training set of known miRNAs) that also does not rely upon cross-species conservation filtering to obtain reasonable results. Thus, this algorithm allows discovery of miRNA binding sites that may not be present even in closely related species. Furthermore, *Rna22* does not make use of seed matches , the pattern approach has no need to know the identity of the targeting miRNA in order to find putative binding sites – permitting the identification of binding sites for miRNAs still unknown and thereby discovery of new miRNAs in a bottom-up approach.

GenMiR++ (generative model for miRNA regulation) (Huang et al. 2007) is a Bayesian data analysis algorithm. It has been used to identify a miRNA-target mRNA regulation network based on RNA expression data from 88 tissues and cell types, sequence complementarity and comparative genomics data. This method was the first attempt to infer miRNA targets based on paired expression profiles of miRNAs and mRNAs.

1.5.2 The problem with predictions – why the present work is necessary

The difficulty of miRNA target finding solely based on computational analysis is evident by the facts that the number of experimentally confirmed heteroduplexes still remains small in respect of the extended effort and that predictions made by different algorithms generally have rather little overlap, sometimes it's even null (Sethupathy et al. 2006).

Animal miRNA target sites are small (causing many hits in the genome, not only in 3' UTRs) and on top of that in the majority of cases there is only partial complementarity between the two RNAs caused by gaps, mismatches and G:U base pairs in many positions – what is clearly harder to model than perfect matches. These facts induce that even small differences in the prediction algorithms can cause a great diversity in results. Apart from those differences it is also not trivial to annotate genes and define their 3' UTR sequences. Despite the collection and annotation of many full length cDNAs, different data sets differ considerably and moreover they often ignore isoforms – like alternative adenylation signals (typically missing in current annotations) where 3' UTRs are composed of a constitutive and an alternative part of varying length, where alternative regions are usually longer by ~2 fold, have higher AU content, and contain more cis elements (Ji et al. 2009). Hence, the same algorithm applied on different data sets (meaning data bases) expectedly generates different results. For example, using the same algorithm on the human RefSeq data set of 3' UTRs and the set of 'known gene' 3' UTRs results in 10-20% variability in predicted relationships (Rajewsky 2006b). This problem is not minor, since over half of all mammalian genes contain multiple polyadenylation sites that lead to different 3' UTRs for a gene, dependent on cell identity (tissue) and even dynamically on cell cycle or state (proliferating or differentiating) and extracellular signals (Ji & Tian n.d.). For human miRNAs this may be especially important, as apparently a large fraction of their target predictions falls into alternative 3' UTR parts of genes with several 3' isoforms (the fraction was 40 % in the cited study but rather old data base versions have been used - 313 human miRNAs from miRBase release 7 and 3' UTR annotations from the *polyA_DB* version 1 based on the hg16 assembly of the human genome) (Majoros & Ohler n.d.). There is a nice and very recent review paper about the eukaryotic 3' end processing machinery and the set of known polyadenylation regulatory factors (*pA factors*), discussing also the interesting interconnections with the transcription and splicing machineries (Millevoi & Vagner 2010).

Another major problem lies in the RNA secondary structure prediction algorithms used so far, such as MFOLD. This popular algorithm (like most others) is based on the *Turner model* that is known to have problems with the correct modeling of loop structures. Its calculations are typically unreliable when the accounted sequence length exceeds a few hundred nucleotides, but, using smaller sequence segments for folding calculations may miss the true in vivo structure (Rajewsky 2006b). However, also in the field of thermodynamic modeling and structure prediction progress is being made. Recently, Aalberts and Nandagopal presented a model based on a two-length freely jointed chain theory (variations of this sort of models are popular in biomechanics and polymer chemistry) that shows significant improvement in accuracy of folding predictions (Aalberts & Nandagopal 2010). Nevertheless, it remains to say that all structure prediction algorithms do not consider mRNA binding factors (proteins) that are known to modulate the mRNA secondary structure.

Anyway, less reliable predictions due to deficient biophysical models are better than no predictions at all. An obvious approach to deal with diverse prediction results is the use of intersections from several algorithms to increase specificity – for the price of losing some sensitivity. Thus, it is common

practice to use multiple algorithms to make more reliable predictions about a particular gene or miRNA. But of course, the convergence of algorithms does not really *prove* the reliability of their predictions, especially when they rely on closely related models and filters. Another strategy to enhance prediction quality by decreasing the false positive rate is the additional use of combined mRNA and miRNA expression profiles obtained from high-throughput methods like *microarray* technology (Huang et al. 2007, Cheng & Li n.d.). There are several methodological approaches to profile (detect and quantify) RNAs – from northern blotting to deep sequencing. These methods and the special challenges of miRNA expression profiling (e.g. their short length or family members differing in only one nucleotide) go beyond the scope, but are discussed in detail in (Benes & Castoldi 2010) and (Kong et al. 2009).

1.6 The in-house prediction data set

The in-house conducted study with the aim two unveil miRNA-mRNA interaction networks related to adipogenesis and obesity (Opriessnig 2008) was designed according to a typical workflow, combining intersection sets of target predictions made by the algorithms *TargetScanS*, *RNA22* and *miRanda* (*TargetScanS* and *miRanda* are the most commonly ones used besides *PicTar*, *RNA22* was chosen because it does not rely on conservation) with paired expression profiles of miRNAs and mRNAs arising from microarray analysis – combining with paired profiles means with correlation analysis results of the mRNA and miRNA expression data sets originating from the same samples. A relevant extension in the conducted study with respect to previous workflows was the integration of *in vitro* data from *human multipotent adipose-derived stem* (hMADS) cells during adipocyte differentiation and *in vivo* data from patient fat tissue samples of the subcutaneous and visceral compartment, thereby providing miRNA/mRNA pair candidates that are more likely to have true regulatory roles *in vivo*.

1.7 Gene reporters – assaying whether the prediction works out

Genetic reporters are commonly used in cell biology to study gene expression and other cellular events, like receptor activity, signal transduction, mRNA processing, protein folding and protein:protein interactions. The two most prominent representatives are based on the genes coding for the *green fluorescence protein* (GFP) and the enzyme family of *luciferases* (*luc*), respectively. GFP and *luc* outperform other reporters in terms of sensitivity and also handling, as both enable light emitting reactions allowing convenient measurement of expression by optical detectors like FACS (fluorescence activated cell sorting) systems and luminometers, or fluorescence microscopes for imaging purposes. However, for certain applications also reporters bearing the DNA sequences for the enzymes *Chloramphenicol acetyltransferase* (CAT; one of the first reporters used in mammals) (Smale 2010b) or *β-Galactosidase* (β-gal) (Smale 2010a) are still in use (e.g. blue-white screening with β-gal). Recently there have been very interesting developments like reporters based on the iron-binding protein *ferritin* which allow the detection of gene expression by magnetic resonance imaging (MRI), enabling dynamic imaging of gene expression and cell migration *in vivo* (B. Cohen et al. 2009), (Gilad et al. 2008).

1.7.1 Bioluminescence

Bioluminescence is the production and emission of cold light by living organisms for communicative purposes. It is a form of chemiluminescence where energy is released in the form of photons (light) by a chemical reaction. This chemical reaction must be sufficiently energetic to produce an excited singlet state molecule that will generate a visible photon when it relaxes back down into its ground state (in contrast to *phosphorescence* and *fluorescence* where sufficiently energetic photons have to be *absorbed*). Oxidation reactions involving molecular oxygen can provide enough energy, what may explain why the primary mechanism in bioluminescent reactions involves the breakdown of a peroxide bond.

The majority of bioluminescent organisms reside in the ocean, indeed, some 80 % of the more than 700 genera (a genus comprises one or several species) known to contain luminous species are marine. Bioluminescence can be found in most of the major marine phyla from bacteria to fish, with comb jellies (jellyfish) having the highest proportion of luminescent species. Photon fluxes produced by these animals span a range from about 10^3 photons per second for a single bioluminescent bacterium to more than 10^{12} for some krill and fish, where the luminescent chemicals are either released directly into the water or retained within cells called *photocytes*. As most bioluminescence has evolved in the ocean, the major part of observed emission spectra is blue, centered in the proximity of 475 nm – the wavelength that travels farthest through seawater. Green being the next most common naturally occurring luminescence color is more often found in shallow, costal species, maybe because increased turbidity of the water scatters blue light and favors the transmission of longer wavelengths. Green or yellow-green is also the common color produced by non-marine species such as fireflies. Indeed, bioluminescent emission spectra extend over the full visible range, but emission of violet, yellow, orange or red color occurs only rarely, and in most of these cases the chemistry behind is still unknown (Widder 2010).

1.7.2 Luciferases and Luciferins

Luciferase (from Latin *Lucifer* derived from *lucem ferre* [lux, lucis = light; ferre = to bring, to bear], meaning light-bearer or light-bringer) is a generic term for enzymes catalyzing chemical reactions causing visible light emission by living organisms (bioluminescence). These enzymes' substrates are a class of light-emitting biological pigments generically termed *luciferins* – thus, this term refers to any molecule utilized by a *luciferase* or a *photoprotein* (see below) to produce light. As there are many different bioluminescent systems and light-producing chemistries, respectively – it is estimated that bioluminescence has evolved independently at least 40 times – there is a need for taxon prefixes to the generic terms, like *Renilla* luciferase (*Rluc*), for instance.

The ability to emit light in association with their substrates and therefore being measured externally to the milieu they reside in has made luciferases an important research tool over the last two decades. Additional major advantages of these enzymes are the very high sensitivity of measurement, the broad linearity range (light output in dependence of enzyme, substrate or cofactor concentration keeping the other two in excess), and the negligible background (virtually no problem with autofluoresence). The two main classes utilized are the beetle (e.g. from the firefly *Photinus pyralis*) and coelenterazine (e.g. from the coral *Renilla reniformis* and the planktonic copepod *Gaussia princeps*) luciferases. As with GFP (see below) only optimized engineered variants

with enhanced light output and stability, and increased expression efficiency are used (Fan & Wood 2007), (Loening et al. 2006). Beside utilization as reporter genes in the common sense luciferases are also incorporated into reporter applications such as resonance energy transfer based sensors (De et al. 2007), (Hoshino et al. 2007) (see **1.7.3** GFP and its relation to luciferases), split reporter complementation systems (Paulmurugan & Gambhir 2003), imaging probes based on luciferases fused to engineered antibodies (Venisnik et al. 2006), or self-illuminating quantum dots by attaching luciferases as internal light source (So et al. 2006) (Loening et al. 2007).

1.7.2.1 Firefly luciferase

Firefly luciferase (*Fluc, FL*) derived from the North American firefly *Photinus pyralis* (EC 1.13.12.7; **Fig. I12**), is an enzyme that catalyzes *D-luciferin* to the electronically excited (optically active) metabolite *oxyluciferin* (Nakatsu et al. 2006). In this process also Mg^{2+}, ATP, and molecular oxygen are necessary, thus the complete reaction has the form:

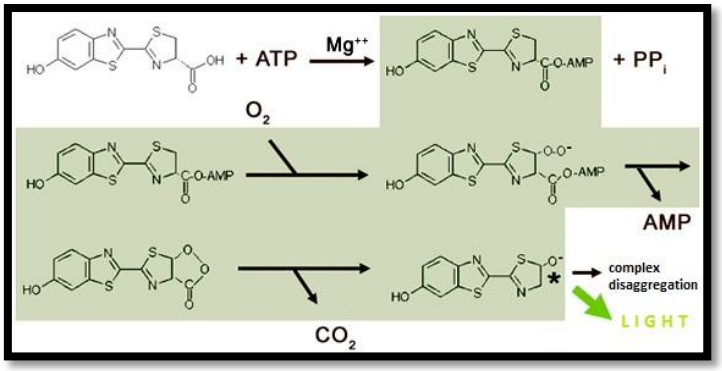


Fig. I10 Chemical reactions catalyzed by firefly luciferase. *Fluc* action is a multistep process, where *D-luciferin* is first activated to *luciferyl adenylate* (releasing *pyrophosphate;* PP_i), then oxidized to *peroxyluciferyl adenylate*, and after release of adenosine monophosphate (AMP) decarboxylated (CO_2 is released) to excited (*) *oxyluciferin*. The transition to its energetic ground state releases the photon (~ 560 nm). During the grey shaded steps the substrate is bound to *Fluc*, final energy state transition induces complex disaggregation. (adapted from MPP 2006)

1.7.2.2 Renilla luciferase

Renilla luciferase (*Rluc, RL*) originating from *Renilla Reniformis* (EC 1.13.12.5; GI 1246926; **Fig. I12**) is a 37 kDa cofactor-less, single subunit enzyme that has a characteristic a/b-hydrolase fold sequence at its core. It catalyses the degradation of its luciferin substrate *Coelenterazine* (CID 2830) in the presence of molecular oxygen resulting in the reaction product *Coelenteramide*, carbon dioxide, and a photon of light (**Fig. I11**) (Loening et al. 2006, Woo & von Arnim n.d.).

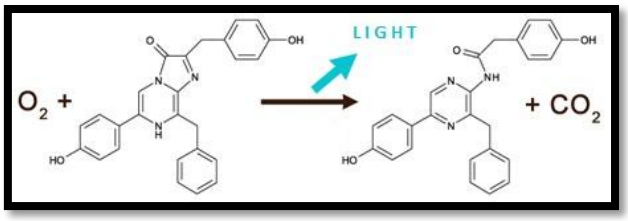


Fig. I11 Chemical reaction catalyzed by *Renilla* luciferase. Compared to firefly, *Renilla* luciferase reaction is a little cheaper. The substrate *Coelenterazine* is directly converted to *Coelenteramide* in an oxidative decarboxylation (with oxygen consumption and CO_2 release). Again the transition from an excited energy state to ground state in *Coelenteramide* releases a photon (~ 480 nm) and destabilizes the complex.

adapted from (MPP 2006)

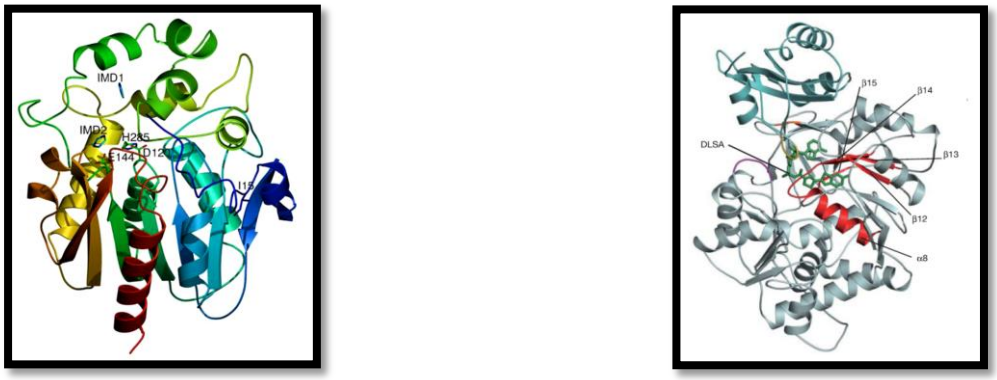


Fig. I12 Ribbon diagrams of *Renilla* and firefly luciferases. (***left***) a variant of *Renilla* luciferase (*Rluc*); The presumptive catalytic triad consists of a *glutamate* (E144), an *aspartate* (D120) and a *histidine* (H285) residue. Also two *imidazole* molecules (known to enhance catalytic activity) located in the catalytic pocket are marked (Loening et al. 2007); (***right***) a variant of firefly luciferase (*Fluc*). The secondary structures for the luciferin-binding site are colored in red; DLSA (in green) is a luciferyl adenylate analogue (Nakatsu et al. 2006);

1.7.3 Green fluorescent protein – and its relation to luciferases

Fig. I13 Structure of the *green fluorescent protein* (GFP)

GFP is composed of 238 amino acids (~ 27 kDa) forming a cylindrical shape consisting of 11 strands of β-sheet (green) forming the walls of a barrel, an α-helix inside containing the chromophore (a Ser-Tyr-Gly sequence) and short helical segments on the ends (blue) capping the top and bottom of the barrel. This so called β-can structure protects the fluorophor in the center (light green) from quenching and photobleaching (F. Yang et al. 1996).

Proteins containing highly visible and efficient fluorophors are naturally produced in several coelenterates (marine organisms) like the jellyfish *Aequorea victoria* (from which the first GFP variant

was isolated) and the sea pansy *Renilla reniformis* – the only two representatives being well characterized. Although many organisms have similar green fluorescent proteins (forming the GFP-like superfamily), the term *GFP* (without additional information about the species) actually refers to the *Aequorea victoria* protein, which is the only one commonly used to date, however, solely in the form of optimized engineered variants with increased fluorescence, photostability, diffusibility and 37° folding efficiency (e.g. *eGFP*; *enhanced*) or alterations in the excitation and emission spectra (e.g. *BYP*, *RYP*, *YFP*, *CFP*; blue, red, yellow and cyan fluorescent), for instance. *Renilla* GFP (with a single excitation peak at 498 nm and an emission peak at 505 nm) and *Aequorea* GFP (with excitation peaks at 395 nm and 475 nm and an emission peak at 505 nm) apparently have the same core chromophore, anyway, *Renilla* GFP (RrGFP) actually has a much higher extinction coefficient and resistance to pH-induced conformational changes and denaturation, caused by an alternative protein-structure. Unfortunately, the *Aequorea* gene is the only one that has been cloned and further developed. Indeed, nature would provide a variety of proteins showing fluorescence, but the biochemistry behind is in most cases completely unknown and it is questionable whether there will be any interest and/or funding to reveal this information. But, there are some interesting approaches for developments of alternative purpose (meaning non-visible light) fluorescent proteins like infrared fluorescent proteins (IFPs, with excitation/emission maxima of 684/708 nm) (Shu et al. 2009). Because in contrast to GFP their wavelengths penetrate tissue well, IFPs are suitable for whole-body imaging.

Nonetheless, the value of GFP is undiminished high (Nobel Price 2008 for its discovery and development). It has become well established as a marker of gene expression, protein localization and dynamics and several protein properties like protein targeting and protein:protein interactions in cell culture (living and fixed) and (with size limits due to absorption) even living organism approaches. This success is mainly caused by its ability to form an internal chromophore without requiring cofactors, enzymes or substrates specific to the originating organisms (besides excitation energy it only needs molecular oxygen [for maturation] to work), the very high stability of the protein (the folding process is temperature sensitive, but once matured properly at lower temperatures GFP is stable and fluorescent at temperatures up to at least 65°C), the apparently low or even missing cytotoxicity (though, high-level expression can be deleterious due to hydrogen peroxide (H_2O_2) release during maturation) and the fact that it allows for tagging on either C- or N-terminus without loss of function of both the GFP and the fused protein (at least in many cases). Unlike luciferases, GFP is not an enzyme that catalyzes an indefinite number of substrate molecules producing a photon per reaction. As fluorescent protein it is generally a secondary light emitter, relying *in vivo* on radiationless energy transfer processes (FRET, fluorescence resonance energy transfer) to produce its green light, thereby shifting the color of the emission to longer wavelengths. In technical use it is excited by irradiation, of course. The functional (exciting) partner in *A. victoria* is *aequorin*, a chemiluminescent protein (*photoprotein*) emitting blue photons when interacting with Ca^{2+} ions (peak emission near 470 nm) and thereby providing the excitation energy for GFP, by contrast in *R. reniformis* GFP is functionally coupled with a luciferase (*Renilla* luciferase) which also emits in the blue range (peak emission at ~ 480 nm) (Chalfie 1995, Tsien 1998).

Photoproteins

Luciferases are one of two classes of bioluminescent proteins that are separated by their relation to luciferin. Where luciferases are usual enzymes (reads, following usual enzyme kinetics) requiring a luciferin as substrate to produce light (see above), the second class, called *photoproteins*, comprises stable complexes of a luciferin (as prosthetic group) and an apoprotein working like a single turnover

enzyme. Photoproteins display luminescence when an additional factor (cofactor) causes a conformational change leading to an intramolecular reaction resulting in an excited state of the luciferin (oxiluciferin) and the moldering of the complex – the relaxation process taking place in the oxiluciferin molecule releases the photon. The luciferin is consumed in this reaction, but the cofactor and apoprotein remain functional and can reenter the process. Indeed, this classification may be unnecessary as photoproteins could also be thought of as very stable enzyme-substrate (luciferase-luciferin) complexes 'waiting for' the cofactor. But, according to this classification *Aequorea* GFP's energy donor *Aequorin* is a *photoprotein* and Ca^{2+} its cofactor, for instance (Chalfie 1995, Tsien 1998).

1.8 Luciferase assay

Generally, in luciferase assays (*luc* assays; LAs) a limiting component of a luciferase-catalyzed reaction is coupled to a variable parameter of interest, while holding the other components non-limiting. Thus, light output varies with the parameter of interest (Cali et al. 2008). A very common use of luciferase-luciferin reactions is the measurement of luciferase enzyme levels depending on transcriptional or translational factors. This can happen in the context of gene expression experiments, where luciferase cDNA expression is placed under control of the regulatory elements under investigation (promoter studies; transcription factor studies), or transcript interaction experiments, where luciferase cDNA is fused to the 3' UTR of a gene (miRNA studies), for instance. Luciferase activity (of firefly) can also be measured as a function of ATP concentration in order to track or exploit a particular chemical reaction, where ATP is supplied or consumed by this reaction – *pyrosequencing* is an important example (Ronaghi et al. 1996). Another important application of LAs is the monitoring of ATP concentration as marker of cell viability or cell death (Sykes & Avery 2009). The luciferin concentration may also be used to gain information about certain reactions. Investigations on drug effects on enzyme activities (e.g. effects on drug-metabolizing CYP and monoamine oxidase enzymes) make use of *luc* assays in that way. In these approaches the enzymes of interest convert inactive derivatives of luciferins to active forms which are detected in a second reaction with luciferases (Cali u. a. 2008).

1.9 Transfection

Transfection is the process of intentionally introducing nucleic acids into eukaryotic cells by nonviral methods (viral methods are referred to as *Transduction*). Although uptake of naked DNA into untreated cells is possible and exploited in some *in vivo* approaches (Wolff & Budker 2005), in cell culture applications enhancing agents or cell membrane manipulations are commonly used to gain strongly increased efficiency (typically for the price of reduced cell viability). An *ideal* nucleic acid delivery method has to meet three major criteria: (1) it must efficiently bring DNA into the cell's nucleus (and RNA into the cytoplasm) and release it there, (2) it must protect the nucleic acids against degradation by nucleases and other enzymes especially during their passage of lysosomes, and (3) the method itself must be non-toxic to the cells. Although viral vectors are effective carriers and would meet the first two criteria, they are not the standard method because of several drawbacks. Usually viruses have to be inactivated in order to eliminate their pathogenic properties. The chance of reversion to a pathogenic virus is small but exists and moreover the use of viruses *in vivo* presents the problem of immunogenicity. In comparison to other methods, the use of viral vectors is typically more expensive

and requires special equipment in order to ensure safety of applicant and environment. Additionally, the virus envelope has a definitive volume and therefore can only deliver limited size DNA. Nonviral delivery systems can overcome most of the problems with viral vectors. The biggest advantages of nonviral transfection reagents are lower (but not absent) immune responses and easier application procedures. Most nonviral gene delivery systems are synthetic materials that can be classified in two groups: cationic lipids and cationic polymers. In both cases amino groups provide the required positive charge (at physiological pH) for DNA packing. In fact, the ability to condense DNA into small non-negative particles is the most important requirement for nonviral gene carriers (Promega T, Fermentas, US Pat. 20100041739, Howell et al. 2003).

1.9.1 Transfection methods

All transfection methods mentioned below have in common that they eliminate the issue of introducing negatively charged nucleic acid molecules (the relevant charge is caused by the phosphate backbones of DNA and RNA) into cells with an also negatively charged membrane. Positively charged chemicals like *calcium phosphate* and cationic lipid or cationic polymer-based reagents coat the nucleic acid molecules and thereby neutralize the negative charge or even create an overall positive charged complex. This net charge compensation or inversion allows the *nucleic acid:transfection reagent* complex closer association with the cell membrane and subsequent crossing by endocytosis or fusion processes.

Calcium phosphate

Calcium phosphate (CaP) co-precipitation is widely used because the components are easily available and inexpensive, and it is effective with many cell lines. The CaP:nucleic acids precipitates (see **3.1.1 Calcium Phosphate Transfection**) are taken up by cells via endocytosis. A benefit of this method is that calcium phosphate appears to provide protection against intracellular and serum nucleases, drawbacks are its sensitivity to nucleic acids amounts and generally its being prone to variability, respectively. In addition, even very small changes in pH (± 0.1) can compromise transfection efficiency, and the method is not suited for *in vivo* transfers (Promega T, Jordan & Wurm 2004, Chowdhury et al. 2004).

Cationic lipids

Cationic liposomes (CLs) are among the prevalent synthetic carriers for nucleic acids (NAs) currently used in biological research as they are one of the most powerful non-viral vectors available. CL-based carriers have the potential of transferring large pieces of DNA of up to 1 million base-pairs into cells. CL-DNA complexes primarily form a multilayered sandwich structure with DNA layered between the cationic lipids (lamellar complexes). On rare occasions, an inverted hexagonal structure with single DNA strands encapsulated in lipid tubules is observed. The membrane cationic charge density has been identified as a key parameter governing the transfection efficiency of lamellar CL-DNA complexes. The current limiting factor to transfection by cationic lipid vectors appears to be the tight association of a fraction of the delivered exogenous DNA with cationic molecules (also cellular),

which may prevent optimal transcriptional activity (Zabner et al. 1995, Ewert et al. 2004, Ewert et al. 2005, Ahmad et al. 2005, Ewert et al. 2008). Stated circumstances for DNA are likely to also apply to RNA molecules (at least dsRNAs such as siRNAs or microRNA mimics), where cationic molecules may interfere with expected regulatory functions.

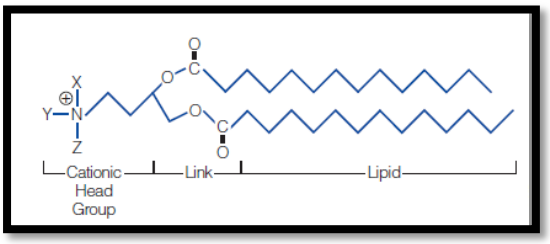


Fig. I14 General structure of a synthetic cationic lipid. X, Y and Z represent possible chemical moieties, which can differ depending on the specific lipid. The cationic head group associates with negatively charged phosphates in the nucleic acid. Often the cationic lipid is mixed with a neutral lipid (such as DOPE, a fusogenic lipid) which enhances the transfer ability (Promega T).

Liposome-mediated delivery offers advantages such as relatively high efficiency of gene transfer (but not necessary transcription), ability to transfect cell types that are resistant to CaP transfection, *in vitro* and *in vivo* applications, delivery of DNA molecules of a large range of sizes (oligonucleotides to yeast artificial chromosomes), delivery of RNA, and delivery of proteins (Promega T and references therein). The main drawback is the fact that lipids generally have membrane-like features what is presumably the reason for their apparent toxicity for cells.

Cationic polymers

Cationic lipids are well-studied DNA carriers, but recently a lot of attention is given to the cationic polymers, which are able to condense DNA into small particles and initiate cellular uptake via endocytosis. However, the transfection efficiency and toxicity of these polymers varies widely. A quite new (patent granting 02.2010) and very promising example is pHP (poly(2-hydroxypropyleneimine)), a cationic polyalkyleneimine polymer (Fermentas, US Pat. 20100041739).

Elektroporation

The mechanism is based on the use of an electric field pulse to perturb the cell membrane integrity for some milliseconds and provoke the formation of transient pores that allow passage of nucleic acids into the cell. Success strongly depends on fine-tuning and optimization of pulse duration and strength for each cell line.

Microinjection

The use of micropipettes to insert substances in single cells is a simple mechanical process utilizing a needle to penetrate the cell membrane. Obviously this method is not practicable in most usual cell culture applications, but it is often the method of choice in experiments with oocytes or early zygotes.

Biolistic particle delivery

Another very physical method of gene transfer is biolistic particle delivery, also known as particle bombardment (by *particle guns*). The method relies upon high-velocity delivery of nucleic acids on microprojectiles penetrating the cell membrane – even in vivo. But, this approach is relatively costly, and in most applications impractical.

Magnet assisted transfection

This method utilizes magnetic nanoparticles loaded with nucleic acids and a magnetic field to transfect cells *in vitro* (Bertram 2006).

(Promega T)

Cell-penetrating peptides and recent constructs

Synthetic or naturally occurring cell-penetrating peptides (CPP) are rather less utilized *in vitro* but they are considered as potential transport system *in vivo*. For *in vivo* approaches also nanoparticles being able to complex DNA and sRNA and being fused to specific ligands which can target them to specific cells expressing the respective receptor are under research. Similar approaches are so called *immunoliposomes.* That is liposome-based transfection as mentioned above, but with liposomes being linked to antibody fragments (Ziello et al. 2010, Puri et al. 2009).

1.10 Thesis objectives

The previously in-house constructed miRNA-mRNA interaction networks related to adipogenesis and obesity (**chapter 1.6**) are based on bioinformatic predictions, where the rather poor reliability of these methods has been outlined (**chapter 1.5**). Experimental validation of theoretical predictions is generally necessary, but in this particular context it is essential not only to proof predictions but also to be able to differentiate between direct and indirect effects of miRNA-based regulation seen in overexpression and silencing studies.

Thus, the main objective of this thesis is the establishment of the *luciferase assay*, a suitable method to provide evidence for in fact occurring direct interactions on the molecular level. The assay should be optimized for easy handling, low toxicity, high nucleic acid delivery and silencing efficiency. A further aim is the validation of a predicted interaction between the mRNA of nuclear receptor PPARγ and the microRNA miR-27b, where also in-house generated experimental data indicate a functional relation.

The specific aims of this thesis are:

1. Literature research for comparative rating and experience-based choice of assay components
2. Establishment of cell culture handling procedures for the model cell lines
3. RNA isolation and miRNA expression profiling of the model cell lines
4. Establishment of a FACS protocol for determination of transfection efficiency and cell viability based on small sample amounts (96-well format)
5. Optimization of cell culture and transfection conditions with associated tests of transfection efficiency and cell viability using eGFP and 7-AAD in FACS analysis and a luciferase vector in combination with siRNA/miRNA in luminescence measurements.
6. Establishment of a luminescence measurement protocol with respect to sample preparation, sample and substrate amounts, and timings.
7. Proof of functionality by verifying a published result, and functional sRNA dose tests
8. Cloning of a high priority candidate construct (PPARγ 3’ UTR) and subsequent validation

Based on literature research the model cell lines HEK293 and HeLa have been chosen to provide cellular environments (where the decision to establish the assay for a second cell line has been made later on, therefore initial CaP experiments were restricted to HEK293) and the transfection reagents DharmaFECT Duo (Dharmacon), METAFECTENE Pro (Biontex), and *Calcium Phosphate* (Sigma-Aldrich) to act as nucleic acid transfer agents. Furthermore, the microplate luminometer *ORION II* (Berthold) and the luciferase vector *psiCHECK-2* (Promega) have been purchased.

2. Materials

2.1 Overview

2.1.1 Cell culture and transfection

MATERIAL	COMPANY
Dulbecco's modified Eagle medium (DMEM) 4.5 g/l glucose, l-glutamine, pyruvate	Gibco, Cat#41966
Fetal bovine serum (FBS)	Pan Biotech, Cat#3302-P210302
L-Glutamine 200 mM	Gibco, Cat#25030-024
Pen-Strep (mixture: 5000 U/ml Penicillin + 5000 µg/ml Streptomycin	Cambrex, Cat#DE17-603E
Normocin (50 mg/ml)	InvivoGen, Cat#ant-nr-2
Phosphate buffered saline (PBS) 1x pH7.4	Gibco, Cat#10010
Trypan blue staining solution	
Trypsin-EDTA 0.5% Trypsin 5.3 mM EDTA.4 Na	Gibco, Cat#
100 mm cell culture dish	Greiner, Cat#664160
150 mm cell culture dish	Greiner, Cat#639160
24-well plate	Corning, Cat#3524
96-well plate	Greiner, Cat#655180
DharmaFECT Duo	Dharmacon, Cat#T-2010
METAFECTENE Pro	Biontex, Cat#T040
Calcium Phosphate Kit	Sigma-Aldrich, Cat#CAPHOS-1KT
miRIDIAN microRNA Mimics and Inhibitors	Dharmacon
miRIDIAN microRNA Mimic Negative Control #1 (miR-NTC)	Dharmacon, Cat#CN-001000-01
siRNA (*siGENOME* Non-Targeting siRNA Pool #2)	Dharmacon, Cat#001206-14
psiCHECK-2 luciferase vector	Promega, Cat#C8021
pEGFP-C1	Clontech, Cat#6084-1
pBluescript II KS-	Stratagene, Cat#212208
Trypan blue staining solution 0.4%	Sigma, Cat#T8154
Cryo Freezing Container	Nalgene, Cat#5100-0001
DMSO	Sigma-Aldrich, Cat#472301-1L
Aqua bidestillata sterilis (Fresenius water), ddH_2O	Fresenius Kabi, Cat#0698961
RNase-free water	Macherey-Nagel
10 µl filtered pipette tips	BioPoint Scientific, Cat#311-4050
100 µl filtered pipette tips	BioPoint Scientific, Cat#361-4050
1000 µl filtered pipette tips	BioPoint Scientific, Cat#342-4050
0.2 ml PCR tubes	Eppendorf, Cat#0030124.332
1.5 ml microcentrifuge tubes	Eppendorf, Cat#0030125.150
2.0 ml microcentrifuge tubes	Eppendorf, Cat#0030123.344

2.1.2 miRNA expression profiling

MATERIAL	COMPANY
TRIzol Reagent	Invitrogen, Cat.#15596-018
Chloroform	Sigma, Cat.#C-2432
Isopropanol	Merck, Cat.#100995
Ethanol	Merck, Cat.#1.08543.0250
ddH_2O	Fresenius Kabi, Cat#0698961
RNase free water (DEPC treated)	Carl Roth, Cat.#T143.3
RNA 6000 Ladder	Agilent
RNA 6000 Nano Dye Concentrate	Agilent
RNA 6000 Nano Marker	Agilent
RNA 6000 Nano Gel Matrix	Agilent
RNase AWAY	Molecular BioProducts, Cat#7003
miRCURY LNA microRNA Array Power labeling kit	Exiqon, Cat#208031
20x SSC	Exiqon, Ambion
10% SDS	Exiqon, Ambion
BSA	Sigma-Aldrich
99.5% formamide	Sigma-Aldrich
2x hybridization buffer	Exiqon
1x hybridization buffer	99.5% formamide (500 µl) + 20x SSC (250 µl) + 10% SDS (10 µl) + MilliQ water (240 µl)
Prehybridization buffer	20x SSC (125 ml) + 10% SDS (5 ml) + BSA (5 g) + MilliQ water (370 ml)
Wash buffer I	20x SSC (100 ml) + 10% SDS (10 ml) + MilliQ water (890 ml)
Wash buffer II	20x SSC (50 ml) + MilliQ water (950 ml)
Wash buffer III	20x SSC (25 ml) + MilliQ water (975 ml)

2.1.3 Cloning, mutagenesis and DNA preparations

MATERIAL	COMPANY
QIAquick PCR Purification Kit	Qiagen, Cat#28106
peqGOLD Gel Extraction Kit	peqlab, Cat#12-2501-02
T4 DNA Ligase	Invitrogen, Cat#15224041
HiFi Taq Polymerase & Buffer	Fermentas, Cat#F-530L
IPTG	Invitrogen, Cat#15529019
GeneRuler 1 kb DNA Ladder	Fermentas, Cat#SM1333
PureLink Quick Miniprep Kit	Invitrogen, Cat#K2100
peqGOLD Universal Agarose	peqlab, Cat#35-1020
DH5-α chemically competent *E.coli*	Invitrogen, Cat#18265017
QIAprep Spin Miniprep Kit	Qiagen, Cat#27106
Not I restriction enzyme	Promega, Cat#R6431
Xho I restriction enzyme	Promega, Cat#R6161
Buffer D	Promega, Cat#R9921
QuikChange Site-Directed Mutagenesis Kit	Stratagene, Cat#210519

2.1.4 FACS measurements

MATERIAL	COMPANY
Phosphate buffered saline (PBS)	Promega, Cat#E1980
7-amino-actinomycin D (7-AAD)	BD Pharmingen, Cat#559925
FACS Flow	BD, Cat#342003
FACS Rinse	BD, Cat#340346
FACS Clean	BD, Cat#340345
Falcon tube, 5 ml, Polystyrene	BD, Cat#352052

2.1.4 Luminometer measurements

MATERIAL	COMPANY
Dual-Luciferase Reporter Assay System	Promega, Cat#E1980
Passive lysis buffer (PLB)	Promega, Cat#E1941
Phosphate buffered saline (PBS)	Gibco, Cat#10010
96-well Assay Plate flat bottom, non-treated, white polystyrene	Costar, Cat#3912
Gelatin	Sigma, Cat#G9391-100G

2.1.5 Instruments

MATERIAL	COMPANY
ORION II Microplate Luminometer	Berthold
FACSCalibur	Becton Dickinson (BD)
TECAN HS400 (Hybridization Station)	Tecan
Nanodrop ND1000 (Spectrophotometer)	Thermo Scientific
Bioanalyzer 2100	Agilent
Centrifuge 5415D	Eppendorf
Centrifuge 5702	Eppendorf
Thermomixer Compact	Eppendorf
Micro Centrifuge	Carl Roth
High volume refrigerated bench top centrifuge 6K 15	Sigma Laboratory Centrifuges
GenePix 4000B scanner	Axon Instruments
Transsonic T420 (US bath)	Elma
PTC -225 Peltier Thermal Cycler	MJ Research
MS2 minishaker	IKA
Hemocytometer	Neubauer
SSM1 mini orbital shaker	Stuart
LaminAir Model 1.2	Holten
Hotplate	Medite
CKX41 Inverted Light Microscope	Olympus
Water bath	Julaba
Incubator	Binder
Incubator/Shaker 3030	GFL
UV Transilluminator	Biorad

2.2 Cell Lines

2.2.1 HEK293 cells

Human embryonic kidney (HEK) 293 cells (sometimes also referred to as 293 cells; 293 is simply an experiment number) were generated in the early 1970s by transformation of normal kidney cells with *adenovirus 5* DNA, where ~4.5 kb from the left viral arm of the viral genome were incorporated into human chromosome 19 (Louis et al. 1997). These cells are commonly considered as easy to grow and readily to transfect. More recently evidence shows that HEK293 cells have many properties of immature neurons, suggesting that in fact a neuronal lineage cell was transformed in the original kidney culture (Shaw et al. 2002).

2.2.2 HeLa cells

Henrietta Lacks (HeLa) cells are one of the oldest (1951) and most commonly used immortal cell lines derived from cervical cancer cells taken from Henrietta Lacks. Horizontal gene transfer from *human papillomavirus 18* (HPV18) to cervical cells created the HeLa genome. As long as provided a suitable environment these cells can divide an unlimited number of times (due to owning an active version of *telomerase*). HeLa cells have a chromosome number of 82 with 4 copies of chromosome 12 and 3 copies of chromosomes 6, 8, and 17. There are many strains as they continue to evolve in cell culture and accumulate aberrations, respectively (Macville et al. 1999).

2.3 Used transfection reagents

2.3.1 DharmaFECT Duo

According to the manufacturer, DharmaFECT Duo (**DFD**) (Promega, Cat#T-2010) is a liposome-based reagent specially formulated to provide high co-transfection efficiencies of small RNAs (miRNA mimics or inhibitors, and siRNAs) with reporter plasmids, and low toxicity under optimized conditions.

2.3.2 METAFECTENE Pro

According to the manufacturer, METAFECTENE Pro (**MP**) (Biontex, Cat#T040) is a liposome-based reagent providing minimal toxicity, enhanced release of nucleic acids in the cytosol (by lower lipoplex stability within the cell) and increased transgene expression rates per cell relative to former reagent generations.

2.3.3 Calcium phosphate

The Calcium phosphate (**CaP**) transfection kit (Sigma, Cat#CAPHOS-1KT) utilized is based on the typical standard components 2.5 M $CaCl_2$ (Cat#2052) and *2xHEPES-buffered saline* pH 7.05 (Cat#H1012).

2.4 Reporter vectors

2.4.1 Luciferase vector

The utilized *psiCHECK-2* vector (Promega, Cat#C8021; AY535007.1) contains a synthetic version of the *Renilla* luciferase (*hRluc*) reporter gene for monitoring RNAi activity. Several restriction sites are included 3' of the luciferase translational stop codon, allowing creation of transcriptional fusions between the sequence of interest and the *Renilla* luciferase reporter gene. Because of the presence of a stop codon in-frame with the *hRluc* open reading frame, no fusion protein is produced. The vector contains an additional reporter gene, a synthetic firefly luciferase (*hluc+*) allowing for normalization of transfection efficiency and cell number without the need of the usual reference plasmid cotransfection. For a vector map see **Fig. A13** (Appendix).

2.4.2 eGFP vector

The utilized *pEGFP-C1* vector (Clontech, Cat#6084-1; kindly provided by *Anne Krogsdam*, IGB) for FACS analyzes encodes a red-shifted variant of wild-type GFP which has been optimized for brighter fluorescence and higher expression in mammalian cells (excitation maximum 488 nm; emission maximum 507 nm). The vector bears a CMV promoter, an SV40 origin, and a SV40 polyA signal.

2.5 microRNA Mimics and Inhibitors

For gain-of-function studies (also termed *overexpression studies*) and loss-of-function studies (or *silencing studies*) chemically synthesized oligonucleotides are available from different manufacturers.

The used *miRIDIAN microRNA Mimics* and *Inhibitors* (Dharmacon) are designed based on known or predicted miRNA sequences in miRBase. *microRNA Mimics* are double-stranded RNA oligonucleotides designed to mimic the function of endogenous mature miRNAs. They are chemically enhanced to preferentially program RISC with the active strand. *microRNA Inhibitors* on the other hand are single-stranded RNA oligonucleotides designed to inhibit the function of endogenous miRNAs by hybridization. They are also chemically enhanced to improve efficiency and biological stability. For both, mimics and inhibitors negative controls are available. These control RNAs are based on *C. elegans* miRNAs and have been confirmed to have minimal sequence identity with miRNAs in human, mouse and rat (Dharmacon).

2.6 7-AAD

7-amino-actinomycin D (7-AAD) (BD Pharmingen, Cat#559925) is a fluorescent agent that intercalates between cytosine and guanine bases of double-stranded nucleic acids. It is excluded by viable cells but can penetrate cell membranes of dying or dead cells, thus it can be used to define dead (7-AAD bright), apoptotic (7-AAD dim), and live (7-AAD negative) cell populations either quantitatively by flow cytometry (FACS) or qualitatively by fluorescence microscopy (Schmid et al. 1992, Philpott et al. 1996).

2.7 Dual-Luciferase Reporter Assay System

The utilized *Dual-Luciferase Reporter Assay System kit* (Promega, Cat#E1980) allows the sequential measurement of both luciferase activities from a single sample. The kit provides substrates and appropriate buffers for both luciferases and a *passive lysis buffer* (PLB) for preparation of cell lysates. The firefly luciferase reporter is measured first by adding *Luciferase Assay Reagent II* (LARII; the substrate for firefly luciferase diluted in buffer) to the cell lysate. After quantifying luminescence an equal amount of *Stop&Glo Reagent* (S&G; the substrate for *Renilla* luciferase diluted in buffer) is added what initiates *Renilla* luciferase action and simultaneously quenches the firefly reaction. Both reagents are optimized to produce stabilized, slowly decaying signals that are linear for both luciferases down to 10^{-20} molecules of *FL* and 10^{-19} molecules of *RL* (Promega, DLR System Technical Manual).

3. Methods

3.1 Cell culture handling of HEK293 and HeLa cells

Based on literature and recommendations of colleagues the same medium conditions could be applied to both cell lines. Standard medium (***Medium I***) consisted of *Dulbecco's Modified Eagle Medium* (DMEM) supplemented with 4.5 g/l D-*glucose*, 4 mM L-*glutamine*, 25 mM HEPES, and *phenol red* (Gibco, Cat#41966), 50U+50µg/ml *Penicillin*/*Streptomycin* (Cambrex, Cat#DE17-603E) and 0.1 mg/ml *Normocin* (InvivoGen, Cat#ant-nr-2); 2 mM L-Glu (L-*glutamine*; Gibco, Cat#25030-024) and 10 % FBS (fetal bovine serum; Pan Biotech, Cat#3302, Lot#P210.302) were added immediately before use. Medium was generally preheated for 15 min in a 37°C water bath before it was applied to cells.

Both cell lines were grown and maintained under normal environmental conditions (37°C, 5 % CO, humidified) in 10 ml of *Medium I* per 100 mm dish (18 ml per 150 mm dish; Greiner) and regularly split (typically 2 to 4 days) by ratios of 1:5 to 1:15 when they reached ~ 70 to 80 % optical confluency. Detachment of cells (both lines are growing adherent) was contrived by trypsinization, where *trypsin* (0.25% *trypsin* in PBS) was preheated to 37°C and 1 ml per 100 mm dish (2 ml per 150 mm dish) was added to the cells after medium removal and two times washing with an appropriate amount of sterile PBS (room temperature). Plates were placed on a 37°C hotplate and incubated for 2 to 5 minutes (in case of HEK293 all cells typically floated upon mechanical assistance for *trypsin* action – some palm strikes – after 2 to 4 min, while with HeLa rather 3 to 5 min were necessary). *Trypsin* was inactivated with 4 times the amount of *Medium I* as soon as possible, wherein cells were

resuspended gently but thoroughly (by pipetting up and down several times using *Corning Stripette* Serological Pipettes). Finally cells were transferred to a new culture plate, which was gently swirled for evenly distribution.

Seeding of cells for subsequent transfection experiments comprised trypsinization as stated and preparation of cell suspensions with appropriate cell concentration (e.g. 200.000 cells per ml for 20.000 cells in 100 µl per well). For that, a hemocytometer and a convenient Excel sheet for calculation of the dilution were used. Routinely *trypan blue* staining has been conducted to check cell viability during counting (10 µl of *trypan blue* were added to 50 µl of cell suspension, mixed by pipetting, and 5 min incubated at room temperature; counted cell number/counted fields x 12.000 yields respective number of cells; implemented in Excel sheet). As antibiotics typically cause cell toxicity during transfection processes and a PBS wash prior to addition of *Medium I* (1 hr before transfection using MP, immediately before transfection using DFD) is rather inconvenient with many wells, cells got centrifuged and resuspended in *Medium I* without antibiotics. 100 µl/500µl of cell suspension was added per well of a 96-well/24-well plate for overnight incubation.

For preparation of cryo stocks, cells were trypsinized, pooled, and counted with a hemocytometer to allow for adjustment of cell density to $1x10^6$ per ml, where cell suspension was centrifuged (300 g for 5 min) and cells were resuspended in DMEM (Gibco, Cat#41966) supplemented with 70% FBS and 10% DMSO (dimethyl sulfoxide; Sigma-Aldrich, Cat#472301-1L). After preparation of 1 ml aliquots in sterile cryovials (Nalgene), tubes were quickly placed in a pre-chilled (-20°C) cryo freezing container (- 1°C/min; Nalgene, Cat#5100-0001) and put to -80°C overnight, and eventually transferred to liquid nitrogen.

To revive cells, cryo stocks were thawed in a 37°C water bath, and immediately after liquefaction the ml of cell suspension was diluted in 6 ml of *Medium I* (without antibiotics). Subsequently, cells were centrifuged at 300 g for 5 min, resuspended in 18 ml of *Medium I* and plated in a 150 mm dish.

3.1 Transfections

Although for many transfection reagents there are fast protocols available where seed and transfection happen at the same day, it is generally recommended to subculture cells ~ 24 hrs before transfection as this improves normal cell metabolism and increases in most cases the likelihood of nucleic acid uptake.

Preparation – Lipid-based Transfections

One day before transfection experiments, cells were seeded in a 96-well tissue culture plate at densities from $1×10^4$ to $3x10^4$ cells per well in a total volume of 100 µl *Medium I*, and incubated at 37°C in a 5% CO_2 incubator until they reached ~ 90 to 100% optical confluency (see **Tab. R1**).

Preparation – Calcium Phosphate Transfections

One day before transfection experiments, cells were seeded in a 96-well or 24-well tissue culture plate at densities from $3×10^3$ to $15x10^3$ cells per well in a total volume of 100 µl or 500 µl of *Medium I*, and incubated at 37°C in a 5% CO_2 incubator until they reached ~ 50 to 60 % optical confluency.

In general, prior to transfection all components were brought to room temperature and sensible stock solutions (nucleic acid solutions, transfection reagents) were mixed gently by agitation. All steps were performed in a laminar flow cell culture hood using sterile technique. '*Pipetting once (or twice), gently and carefully'* included swirling the pipette tip a little bit.

3.1.1 Calcium phosphate transfections

The main drawback of this method is its high sensitivity to the amount of input nucleic acids (NAs) – amplifying pipetting inaccuracies, and the generally high amount of total nucleic acid to obtain good transfection efficiencies. Due to the latter fact *fill DNA* (*pBluescript*, BSC) was used to increase total amounts when necessary.

For CaP transfections the provided kit by Sigma-Aldrich was used according to manufacturer's protocol as far as practicable. Briefly, two hours before transfection medium was replaced with 100 µl (96-well) or 500 µl (24-well) of *Medium I* without antibiotics (and 0 to 10% serum; after initial tests no serum has been added at this point in subsequent experiments). Transfection medium was prepared in 1.5 ml microcentrifuge tubes (Eppendorf) by dilution of ~1 µg nucleic acid (reporter plasmid, siRNA/miRNA and *pBluescript* as neutral carrier plasmid for maintenance of total nucleic acid amount) and 1.8 µl 2.5M $CaCl_2$ in 18 µl *molecular biology water* (Sigma, Cat#W4502) for the 24-well format, or ~235 ng nucleic acid and 0.28 µl 2.5M $CaCl_2$ in 2.8 µl water for the 96-well format – solutions were gently mixed by pipetting. Subsequently, the same total amounts (19.8 µl or 3.1 µl) of 2 x HEPES-Buffered Saline (HeBS, pH7.05) have been added. In case of 24-well the standard mixture procedure could be applied, thus HeBS was added in a second sterile 1.5 ml tube and bubbled using an automatic pipette pump attached to a 5 ml sterile serological pipette (filter plugged). While bubbling the HeBS, $CaCl_2$/NA solution was added dropwise with a sterile pipette tip and eventually the solution was vortexed for 3 seconds. In case of 96-well the HeBS has been pipetted directly into the $CaCl_2$/NA containing tube but without mixing the solutions - mixture was done by vortexing for 5 sec at high speed. In both cases after vortex-mixing the solution was spun down and allowed to sit undisturbed for 20 min at room temperature to form precipitates. After this incubation time precipitates were added to cells and plates were gently agitated to enhance distribution. After an incubation for 4 to 16 hrs (final protocol 16 hrs) under 37°C standard conditions (only in initial experiments after 4 or 16 hrs additional glycerol shocks have been conducted) the transfection medium was replaced with *Medium I* (10% serum) and cells were incubated for further 48 hrs until harvest for analyses.

Glycerol shock

After 4 or 16 hrs transfection medium was aspirated and 100 µl (24-well) of a glycerol solution (1 ml 50% (w/v) sterile *glycerol*, 2.5 ml 2 x HeBS, 1.5 ml molecular biology water) was added. After an incubation of exactly 2 min at room temperature glycerol was aspirated, cells were washed twice with sterile PBS and 500 µl fresh *Medium I* were added. *Glycerol* induces an osmotic shock (see **Fig. A9**, Appendix) and thereby potentially improves expression of transfected plasmids.

3.1.2 METAFECTENE PRO transfections

MP transfections were conducted according to manufacturer's protocol. Briefly, one hour before transfection cells were provided with fresh *Medium I* (without antibiotics) – enhancing proliferation rate. Stock plasmids (~ 50 ng/µl) and siRNA/miRNA (2 µM, 5 µM) solutions were prepared in pH 7.4 buffered ddH_2O (Fresenius Kabi) and RNase-free water (Macherey-Nagel), respectively. Per well ~ 25 to 250 ng of nucleic acids (25 to 250 ng DNA plus 0 to 100 nM sRNA) were diluted in 15 µl PBS, and 0.1 to 1 µl of MP in 10 µl PBS. (As the liposomes in MP show a strong tendency to adhere on walls of typical cell culture vessels PBS *must* be pipetted first to avoid a drop in transfection efficiency, likewise nucleic acids and biomolecules in general tend to adhere to materials typically used.) Solutions were prepared in 1.5 ml tubes (Eppendorf, Cat#) or 0.2 ml tubes (Eppendorf, Cat#), mixed gently by pipetting one time, followed by an incubation of 5 minutes (do not exceed substantially) at room temperature. Subsequently, nucleic acids solutions were added to MP solutions (respect the order) without any mixture procedure and incubated for 20 min (do not exceed 25 min). Immediately after incubation time nucleic acid-lipid complexes have been added to the cells (where dropwise addition and swirling of the plate are not necessary in this small format), and placed in a 37°C 5% CO_2 incubator for 6 hrs (3 to 24 hrs in timing experiment). Thereafter transfection medium was removed, fresh *Medium I* was added and incubation continued until harvest.

3.1.3 DharmaFECT Duo transfections

DFD transfections were conducted according to a slightly modified version of the manufacturer's protocol. Briefly, stock plasmids (~50 ng/µl) and siRNA/miRNA (2 µM, 5 µM) solutions were prepared in pH 7.4 buffered ddH_2O (Fresenius Kabi) and RNase-free water (Macherey-Nagel), respectively. Per well ~25 to 360 ng of nucleic acids (25 to 250 ng DNA plus 0 to 200 nM sRNA) were diluted in 10 µl DMEM, and 0.05 to 0.6 µl of DFD also in 10 µl DMEM. Solutions were prepared in 1.5 ml tubes (Eppendorf) or 0.2 ml tubes (Eppendorf), mixed gently by pipetting one time, followed by an incubation of 5 minutes (do not exceed substantially) at room temperature. Subsequently, DFD solutions were added to nucleic acid solutions, mixed gently by carefully pipetting up and down once and incubated for 20 min (do not exceed 25 min). In the interim medium was removed from the cells and replaced by 80 µl *Medium I* without antibiotics (a medium change one hour prior to transfection as with MP and many other reagents did not show significant positive effects). Finally, immediately after incubation time 20 µl of transfection medium were added (as with MP dropwise manner and swirling of plate is not necessary), and cells were placed in a 37°C 5% CO2 incubator until harvest (removal of transfection medium is not necessary with both cell lines; determined by *trypan blue* staining in 24 hr intervals).

3.2 FACS measurements

3.2.1 Sample preparation

Cells were harvested by trypsinization (50 µl *trypsin* per well|96-well plate; 37°C) for 2 to 4 minutes (HEK293) and 3 to 5 minutes (HeLa), respectively. After this incubation time the culture plate was briefly and very slightly brought into contact with a vortexer to enhance detaching. *Trypsin* was

inactivated with 200 µl DMEM supplemented with 20 % FBS. To singularize cells each cell suspension was gently pipetted up and down 10 times and transferred to 1.5 ml Eppendorf tube. In the next step cell suspensions were centrifuged (300 g for 5 min) and resuspended in 100 µl PBS, 1 µl of 7-AAD was added and cells were incubated for 15 min in the dark at 4°C. Samples have not been centrifuged and washed again to remove 7-AAD, thus they were directly transferred to FACS tubes (BD, Cat#352052) and kept on 4°C in the dark. Due to declining cell membrane integrity over time (substantially starting from ~2.5 hrs) samples have been measured within 2 hrs.

3.2.2 Measurement and data analysis

Measurements were conducted on a FACSCalibur (BD) using *CellQuest Pro* software (BD) for data acquisition and *WinMDI 2.3* (Joseph Trotter; freeware) for subsequent offline data analysis. eGFP emission was measured using a 530r30-nm band pass filter (FL1 channel), and the 7-AAD emission was detected with a 650-nm long pass filter (FL3 channel), whereat the FL3 channel had to be compensated for signal bleeding from eGFP (indirectly per FL2; 10%). Totals of 15.000 cells were measured for each sample. Controls included untreated (untransfected; GFP^-) cells for determination of autofluorescence levels and assessment of sample preparation (7-AAD uptake) and cell suspensions incubated at 50 °C for 30 min (7-AAD positives). Untreated cells were measured in the beginning and in the end of a measurement session to check for 7-AAD uptake in untreated cells due to prolonged periods until measurement. Gating strategies included exclusion of cell fragments (due to pipetting) and doublet discrimination.

3.3 RNA isolation and quality control

RNA isolation with *TRIzol* (Invitrogen, Cat#15596-026) was conducted according to SOP MET023_0 (IGB). Briefly, for harvesting, cells were lysed in 6 ml *TRIzol* per 100 mm dish using a cell scraper to detach them from the surface. Subsequently, lysates were homogenized by vortexing and separated by addition of 0.2 ml *chloroform* (Sigma, Cat#C2432) per 1 ml *TRIzol* followed by shaking for 2 min and further 2 min of incubation at room temperature. Centrifugation at 12.000 g for 17 min at 4°C split the mixture into a lower red phenol-chloroform phase (organic phase), an interphase and the upper colorless aqueous phase – where RNA remains exclusively in the aqueous phase and DNA and protein are in the other two. The aqueous phase was carefully pipetted off, transferred to a fresh tube, and treated with 0.5 ml *isopropanol* (Merck, Cat#100995) per 1 ml *TRIzol* to precipitate the RNA. After 10 min incubation at room temperature the sample was centrifuged at 12.000 g for 20 min at 4°C. Subsequently, the supernatant was carefully removed and the RNA pellet was washed by vortexing with 75% ethanol (1 ml per 1 ml *TRIzol*) and subsequent centrifugation at 7.500 g for 6 min at 4°C. Thereafter the ethanol wash was pipetted out and the RNA pellet was air-dried for some minutes until most ethanol evaporated. Finally RNA was dissolved in 25 µl RNase-free water (DEPC-treated) and incubated for 10 min at 55°C. After concentration and purity was determined total RNA preparation was stored at -80 °C. RNA concentration and purity was assessed by spectrophotometry using a *Nanodrop ND1000* (Thermo Scientific). Measuring 1.5 µl per sample against the appropriate blank (DEPC-treated water) yielded the ratio of absorbance at 260 and 280 nm, where an A260/A280 ratio of 1.8 to 2.1 has been considered as straight. RNA integrity was checked on an Agilent *Bioanalyzer 2100* according to manual.

3.4 miRNA expression profiling

An in-house established microarray production platform was used to generate the oligonucleotide printed chips (prepared chips were kindly provided by Dr. Marcel Scheideler). As solid support serve glass slides coated with a reflective dielectric layer (*Nexterion HiSens E*, Schott) improving signal levels and sensitivity. The spotted miRNA probes are locked nucleic acids (LNAs) (Exiqon, miRCURY LNA microRNA Array ready to spot probe set; Cat#208310), modified RNA oligonucleotides, where the ribose moiety contains an extra methylene bridge connecting the 2' oxygen and 4' carbon. This modification locks the ribose in the 3'-*endo* conformation what alters backbone organization and significantly enhances hybridization properties (melting temperature), causing increased sensitivity and specificity of expression analysis (Kumar et al. 1998, Kaur et al. 2006). Per slide each probe was spotted 4 times (technical replicates) and in addition the whole probe set was spotted twice (local separation) to allow for statistical analysis and to reduce measurement variability (Hackl et al. 2004). The used chips included 1891 capture probes covering all human, mouse and rat microRNAs annotated in miRBase 14.0. In addition, the array contained probes for 385 *miRPlus* human microRNAs - proprietary sequences not found in miRBase (Exiqon).

3.4.1 Sample preparation

Labeling Of microRNA

The labeling of RNA has been done using the miRCURY LNA microRNA Array Power labeling kit (Exiqon, Cat#208031) according to manufacturer's protocol. This labeling system is realized as a 2-step procedure. The first step includes a *calf intestinal alkaline phosphatase* (CIP) for removal of 5' phosphates from terminal of the miRNAs and in the second step a fluorescent label (a single fluorophor per molecule) is attached enzymatically to the 3'-end of the miRNAs in the total RNA sample. This is followed by an enzyme inactivation step after which the sample is ready for hybridization.

Briefly, after thawing all components on ice for 20 min and gentle mixing, 5 µg of total RNA per sample were diluted in 3 µl RNase-free water and a CIP mastermix (per sample: 0.5 µl CIP buffer + 0.5 µl CIP enzyme) was prepared. RNA and CIP solutions were combined in 0.2 ml RNase-free microcentrifuge tubes mixed by pipetting and incubated for 30 min at 37°C and 5 min at 95°C using a PCR cycler. The 95°C enzyme reaction stop was immediately followed by snap cooling on ice for 5 min. In the interim a mastermix containing per sample 3 µl labeling buffer, 2 µl DMSO and 2 µl labeling enzyme was prepared on ice. Subsequently the 4 µl CIP reaction and 1.5 µl of either the fluorophor Hy3 or Hy5 was added. Reagents were gently mixed by vortexing and incubated for 1 h at 16°C (PCR cycler) with heated lid - fluorophors and reaction were protected from light any time. Finally the labeling procedure was stopped by incubation for 15 min at 65°C (PCR cycler) and the sample was left on 4°C until hybridization.

Hybridization

The hybridization procedure was conducted according to SOP MET03_00 (IGB). Briefly, labeled miRNAs were hybridized on an automatic hybridization station (Tecan HS400). This device also automates the final washing steps and performs nitrogen drying. A condition was processed twice using the dye-swap technique for reducing systemic bias (Hackl et al. 2004). Thus, 2 slides were processed in the HEK293 vs. HeLa profiling where once HEK293 was labeled with Hy3 and HeLa with Hy5 and once vice versa.

After loading tubes, liquid channels and hybridization chamber (priming) a pre-wash step using *wash buffer X* (0.2 % SDS) was carried out. Buffers were heated during the whole process to avoid precipitation of solutes. To prevent photobleaching chambers were protected against light. After this initial steps 90 µl of pre-heated (20 min at 65 °C) *pre-hybridization buffer* were injected into the chamber and hybridized for 20 min. This buffer contains BSA (bovine serum albumin) and SDS and is required to prevent unspecific binding of labeled samples on the solid support. Subsequently, a second wash step with *buffer X* was performed. In the interim the labeled RNA samples were combined (Hy5 HEK293/Hy3 HeLa and vice versa), completed with 25 µl RNase-free water and 50 µl 2x hybridization buffer (giving a total volume of 100 µl per slide), and placed in a heating block for 3 min at 90° C for RNA denaturation. Next, 90 µl of this RNA solution were injected per chamber using inverse pipetting to avoid air bubbles and hybridization procedure was automatically performed for the following 16 hrs at 64 °C. Finally, the slide was washed with *wash buffers I, II and III*, followed by nitrogen drying.

3.4.2 Data acquisition and analysis

The microarray slides were scanned using a *GenePix 4000B* microarray scanner and analyzed in the first place with the dedicated software *GenePix Pro 4.1* (Axon Instruments). For grid alignment and adjustment of scanner settings (PMT gain) the spotted probe set included *anchor spots*, capture probes labeled with Hy3 and clearly visible in any case, and several control spots. PMT gains for both channels (λ_1 = 635 nm, red, Hy5; λ_1 = 532 nm, green, Cy3) were also adjusted based on sample probe intensity values, where voltage levels were chosen such that as possible no spot exhibited a very weak intensity value or on the other end ran into PMT saturation. *GenePix Pro 4.1* provides a flag feature option for data filtering relying on the quality of a spot. The classification of a spot is based on shape, intensity, standard deviation, mean and median values. Standard settings were used and spot intensities close to the background intensity or indicating a saturated PMT were flagged and excluded from further analysis.

Subsequent data normalization was conducted with *ArrayNorm* (in-house established; (Pieler et al. 2004)). As initial step data was corrected for the background to eliminate signals originating from non-specific binding of labeled samples on the glass surface or noise from the scanner. Applied was the method of local background subtraction, where the median of the pixel intensities surrounding a spot is subtracted from the spot values. The thereby obtained intensity values are expected to represent unbiased estimators of the true signals. Next, normalization was performed to minimize systematic and random variations in the microarray experiment. The most important systemic variation arises from physical dye properties (light intensity, half-life), the efficiency of dye incorporation (labeling efficiency) and the experimental variability in hybridization and scanner setup

(Hackl et al. 2004). Thus, within-slide normalization (global mean; total intensity normalization) and self normalization (dye swap; replicate filtering) was applied (see (Quackenbush 2002)). Finally, the intensity ratios (R/G corresponding to Hy5/Cy3) were scaled to equal one or zero expressed as $\log_2$ ratio.

For final data visualization (*heatmap*, **Fig. R4**) the in-house established software suit *Genesis* (Sturn et al. 2002) was used.

3.5 Luminometer measurements

3.5.1 Sample and substrate preparation

Except for timing experiments, samples were generally harvested after 48 hrs of incubation (post transfection). Medium was removed and cells were washed once with 30 µl PBS. Next, PBS was replaced with 30 µl of 1x *passive lysis buffer* (PLB; Promega, Cat#E1941) (can be increased for pre-dilution), and the plate was placed on a circular shaker (80 rpm) for 20 min (HEK293) and 30 min (HeLa), respectively. Subsequently the whole plate was either put on ice for subsequently following measurement or to -80°C for long-term storage. Assay reagent for RL (*Stop & Glo*) was prepared freshly from substrate and buffer for each measurement session as recommended, whereas FL reagent (*LARII*) was used from stocks (stored at -80°C) in most cases. Reagents and buffers were generally thawed for 30 min in a water bath kept on room temperature.

3.5.2 Data acquisition and analysis

Luminescence was measured using a *ORION II* microplate luminometer (Berthold) and the dedicated software *Simplicity*. There are several programs available for Dual-Luciferase (DL) measurements allowing for a quick change to a different amount of substrate solution being injected per well (25, 50, and 100 µl program). In any case photon flux integration time is set to 10 s and pre-measurement delay to 2 s, but indeed typical samples would allow to significantly shorten (e.g. certainly down to 5 s) read time. Questions on machine and software handling might be answered by the respective manual.

Simplicity outputs only raw data that is *relative light units* per second (RLU) – in the actual machine setting this corresponds to photons/s. The output file format is XLS, thus subsequent data analysis has to be done in Microsoft *Excel* or the like. Usual steps in *Excel* are calculation of RL/FL ratios, statistical values, statistical significance tests, normalizations and graphical representation.

3.6 Cloning

PCR And Gel Electrophoresis

The human PPARγ2 3' UTR (210 bp) was amplified from cDNA originating from hMADS cells by *polymerase chain reaction* (PCR) using *HiFi Taq polymerase* (Fermentas, Cat#F-530L) and the primers shown in **Tab. A1** (Appendix). All reagents were gently vortexed and briefly centrifuged after thawing and per reaction 40.5 µl ddH_2O (Fresenius Kabi), 5 µl 10 x *HiFi PCR buffer* + MgCl2 (Fermentas, Cat#F-

530L), 1 µl 10 mM dNTP mix, 1 µl 20 µM primer mix, 2 µl template cDNA (100 ng) and 0.5 µl *HiFi PCR enzyme mix* were added in a PCR tube – preparation was done on ice. The samples were placed in a thermo-cycler (PTC-225 Peltier Thermal Cycler, MJ Research) and PCR was run using the temperature profile: 94 °C 5 min; 40 cycles: 94 °C 30 s, 60 °C 30 s, 72 °C, 30 s; 72 °C 10 min. Reaction was kept at 4 °C for short-term storage. Subsequently, *agarose gel electrophoresis* was performed to check for amplification of a correct length product. 24 µl of PCR product were mixed with 12 µl of 3 x loading dye and the whole volume was loaded onto a 1% *agarose* gel. Upon verification of a corrected-sized product (~ 200 bp) the band was cut out of the gel, transferred to a 1.5 ml microcentrifuge tube, weighed and 200 µl of *binding buffer XP2* (peqGOLD Gel Extraction Kit; peqlab, Cat#12-2501-02) per 20 mg gel were added. Next, the tube was incubated at 60 °C for 7 min, where the content was mixed every 2 min. In the following the DNA/agarose solution was loaded onto *HiBind spin columns* and several steps of centrifugation and buffer treatment, and elution were conducted according to manufacturer's protocol to finally yield the extracted DNA in ddH_2O. DNA concentration and purity was determined using the *Nanodrop ND1000* (Thermo Scientific) as mentioned above – with ddH_2O as blank.

Double Restriction Enzyme Digest

For digestion of the insert 50 µl (20 ng/µl) insert solution, 0.6 µl BSA (acetylated, 10 mg/ml; Promega), 6 µl restriction enzyme 10x buffer (*Buffer D*; Promega, Cat#R9921), 2.4 µl ddH_2O and 1 µl restriction enzyme I (*Not I*; Promega, Cat#R6431) *and* 1 µl restriction enzyme II (*Xho I*; Promega, Cat#R6161) were combined in a 1.5 ml microcentrifuge tube. *psiCHECK-2* vector digestion was done combining 1 µl (1 µg/µl) DNA solution, 0.2 µl BSA, 2 µl *Buffer D*, 15.8 µl ddH_2O and 1 µl *Not I and* 1 µl *Xho I*. Also single enzyme control cuts have been performed. Digestion was allowed for 1 h at 37 °C. Subsequently, agarose gel electrophoreses (1% gel) was used to confirm cutting and again *peqlab*'s kit was used to extract the cut vector. The cut insert was purified using the *PCR Purification Kit* from Qiagen (Cat#28106) according to manufacturer's protocol. Concentrations and purity of cut DNAs were determined by photo spectrometry.

Ligation

The ligation mixture containing 50 ng vector and 5x insert was prepared in a 1.5 ml microcentrifuge tube by combining ddH_2O, cut insert and cut vector to a total volume of 20 µl considering the relation (molar ratio): 50 ng vector / size of vector (bp) * size of insert (bp) → weight of insert * 5. Thus, 8.4 ng insert were added. The mixture was incubated for 5 min at 65 °C and then briefly placed on ice. 4 µl 10x *T4 DNA Ligase Buffer* (Invitrogen, Cat#15224041) and 1 µl T4 DNA Ligase (1 Unit) (Invitrogen, Cat#15224041) were added followed by a 2 hr incubation at room temperature. Finally, ligation was put to -20 °C for storage. Relegation controls have been performed.

Transformation

50 µl aliquots of DH5-cells were thawed on ice for each ligation. The tubes containing the thawed ligation were briefly centrifuged and placed on ice. 10 µl per ligation were pipetted directly into the cells and mixed gently by tapping. Subsequently, cells were incubated for 30 min on ice, and then heat shocked for 20 s in a 42 °C water bath without shaking. Next, tubes were placed on ice for 2 min and 300 µl of pre-warmed SOC medium were added. Tubes were placed in an incubator for exactly 1 hour at 37 °C and shook at 225 rpm. Afterwards the whole transformation mixture was spread on a LB agar plate supplemented with 100 µg/ml *Ampicillin* and incubated overnight at 37 °C.

Colony PCR

In a PCR tube 11 µl sterile deionized water, 2 µl 10x *Taq Buffer* + KCl (Fermentas, Cat#EP0402), 1.2 µl 25 mM $MgCl_2$, 1.6 µl 2.5 mM dNTP mix, 2 µl 1 µM primer forward, 2 µl 1 µM primer reverse, 0,2 µl 5 U/µl *Taq DNA polymerase* (Fermentas, Cat#EP0402) and template DNA (a picked colony) were added – the rest of the colony was spread on a LB agar plate. Subsequently, PCR was performed with the temperature profile (98 °C 2min; 35 cycles: 94 °C 30 s, 60 °C 30 s; 72 °C 3 min; 72 °C 10 min; 4 °C ∞). Positive colony PCR clones where determined by gel electrophoreses and respective colonies were picked from the connected freshly streaked LB agar plates and inoculated in 5 ml LB medium containing 100 µg/ml Ampicillin. After overnight (15 hrs) incubation at 37 °C with shaking at 225 rpm a miniprep (*Qiaprep Spin Miniprep Kit*; Qiagen, Cat#27106) was conducted according to manufacturer's protocol. Finally DNA was eluted in ddH_2O and stored at – 20°C.

The usual asymmetrical double restriction step was omitted and correct-sized samples were directly sent to sequencing (Barcode sequencing service, AGOWA). Sample preparation was done according to AGOWA's standards. For sequencing primers see **Tab. A1** (Appendix).

3.7 Mutagenesis

Mutagenesis was conducted using the *QuikChange Site-Directed Mutagenesis Kit* (Stratagene, Cat#210519) according to manufacturer's protocol. Stratagene's kit allows site-specific mutation in double-stranded plasmids eliminating the need for subcloning and for ssDNA rescue. It is possible to insert several point mutations and potentially replace, delete or insert multiple amino acids (in coding regions). The procedure utilizes a dsDNA vector with an insert of interest and two synthetic oligonucleotide primers, both containing the desired mutation (see **Tab. A1**, Appendix). In a PCR run extension of the mutated primers generates a mutated plasmid containing staggered nicks. Treatment with *Dpn I* endonuclease (target sequence: 5'-Gm^6ATC-3') is used to digest the parental (methylated) DNA template and to select for mutation-containing synthesized DNA.

Briefly, mutation primers were designed according to Stratagene's design guidelines using their online tool and NCBI tools and synthesized and PAGE purified by IDT (Integrated DNA Technologies). PCR was prepared and conducted according to protocol using 3.5 min at 68 °C in the elongation step. Subsequent *Dpn I* digestion was done according to protocol with a prolonged incubation of 10 min at 37 °C. Finally, transformation of *XL10 Gold Ultracompetent Cells* was conducted exactly according to manufacturer's protocol. Blue-white screening was carried out to test the efficiency of mutant plasmid generation (> 85 %). Some positive colonies were picked and inoculated in 5 ml LB medium containing 100 µg/ml Ampicillin. After overnight (15 hrs) incubation at 37 °C with shaking at 225 rpm a miniprep (*Qiaprep Spin Miniprep Kit*; Qiagen, Cat#27106) was conducted according to manufacturer's protocol and plasmids were sent to AGOWA for sequencing. Positive plasmids were subsequently used for transformation of DH5-α cells and preparation of glycerol stocks.

3.8 Plasmid preparation and quality control

Of all plasmids used in subsequent transfection experiments glycerol stocks were prepared by adding 0.5 ml of overnight culture (DH5-α cells resuspended in LB Medium without antibiotics) to 0.5 ml of 80 % sterile glycerol. After mixing by inversion (several times) stocks were frozen at – 80°C. Plasmid DNA mini and midi preparations were done using the kits *PureLink HiPure Plasmid DNA Purificatin Kit* (Invitrogen, Cat#K2100) and (*Qiaprep Spin Miniprep Kit*; Qiagen, Cat#27106) according to manufacturers' protocols, where DNA was eluted in ddH_2O (Fresenius Kabi). Spectrophotometry (Nanodrop 1000; Thermo Scientific) was performed to determine DNA concentration and purity (A260/A280 ratio of 1.8 to 2.1) and gel electrophoresis was conducted to check plasmid size, conformation and integrity.

3.9 Data presentation and statistical analysis

Data are presented as mean ± standard deviation. Differences between groups were statistically tested with a *Student's* two-tailed t-test for independent samples with equal variance. In all cases where data have been normalized to a reference condition (all figures with label *relative luciferase activity* and a control condition normalized to 1 or 100; but not normalized viability values in FACS analyses) propagation of uncertainty has been applied – increasing displayed standard deviations. Most experiments conducted with respect to protocol establishment have *not* been biologically (independently) replicated fully due to time and material consumption, but were somewhat checked by repeating a few control conditions. Experiments with biological relevance have been independently replicated (TTP/miR-29a 1 time; PPARγ/miR-27b 2 times) in terms of transfections at different days (different cell passage numbers, transfection media, etc.). In any case at least one technical replicate per condition has been realized.

4. Results

Preamble

Optimizations of cell culture conditions, transfections, the FACS protocol, and the luminescence measurement protocol generated a rather great quantity of experimental data. Reasonably in most cases only essential and representative results are presented to outline the facts decisions on final protocol steps were based on. Especially data from unreplicated small-scale experiments and virtually all initially generated data for the 24-well format are omitted (see exceptions in **Appendix**). This concerns in particular results from *calcium phosphate* transfections where the common disadvantage of low reproducibility (according to literature mainly caused by variation in transfection complex size and shape) has led to an early exclusion of further experiments. Typically these variations are caused by minor changes in the pH of the solutions used for the transfection, and also the manner in which these solutions are combined – thus handling capabilities of the experimenter strongly influence results, especially because the normal manner of mixture (bubbling) is hardly applicable with the low amounts of liquids used in 24-well format and not at all using 96-well format

(assuming amounts for triplicates in both cases), where larger formats did not come into question because of material consumption. Although, the internal control of dual luciferase assays is designed to compensate (among other things) for varying transfection efficiency, stability of transfection results has been an essential point in the decision of which transfection reagent to use eventually. Besides the fact that strongly different amounts of lipoplexes or CaP-nucleic acid precipitates entering cells may lead to different cellular responses (all methods trigger some degree of cellular defense or influence cell signaling (Lonez et al. 2008)) and thus environmental conditions for the transcription/translation/RNAi systems, this emphasis also makes particularly sense as varying transfection success causes varying heights of raw signal (RLU) values in luminometer measurements with related effects on final RLU ratios due to principle of measurement (see **Fig. R31** et seq.) causing increased variation in replicates with according effects (**Fig. R38** et seq.). All transfection reagents used showed low cytotoxicity in mock transfections where they generally cause 7-AAD-detectable negative effects on cell membranes in less than 5% of counted events at the respective highest given concentration (data not shown). The circumstance that transfection complexes are more toxic than reagents (especially lipids) alone is well known, but not understood so far (Biontex).

4.1 Growth curves of HEK293 and HeLa cells

In the context of lipid-based methods RNA transfections are typically conducted at 40-60% *optical confluence* (percentage of growth surface covered with cells), whereas for DNA transfections >80% is recommended for most cell lines and carrier molecules (often 90-100%, depending on the particular transfection reagent). *True confluence* is reached when contact inhibition occurs (meaning inhibition of cell division and cell motility when cells are in close contact with each other). Indeed, for most cell lines optical confluence just indicates the beginning of the log-phase (Biontex). The curves in **Fig. R1** and **Fig. R2** depict the growth dynamics of both cell lines, where the region of optical confluence is highlighted in red. Similar results for optical and true confluence (normalized) were achieved in 24-well format (data not shown).

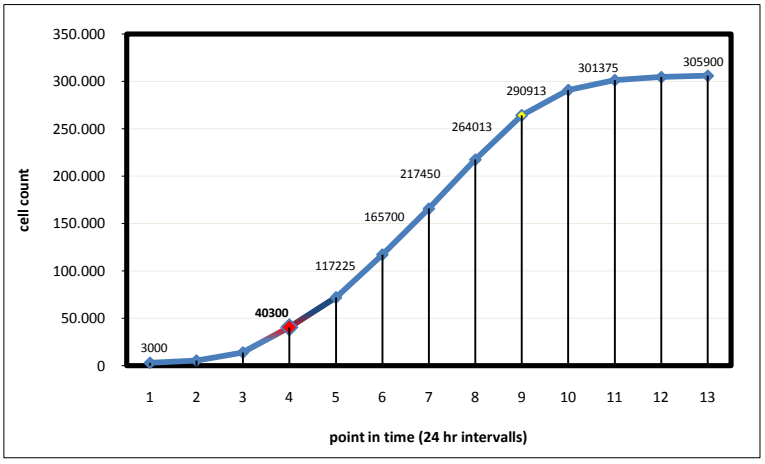


Fig. R1 HEK293 growth curve for the 96-well format. The region where cell density approaches *optical confluence* is highlighted in red.

For the HEK293 cell line (**Fig. R1**) optical confluence corresponds to about 40.000 to 45.000 cells per well (130.000 to 150.000 cells/cm²). True confluence due to contact inhibition indicated by the

transition from log growth phase into plateau phase starts out at approximately 270.000 cells per well (900.000 cells/cm²) and permits a maximal cell number of roughly 300.000 cells per well (1.000.000 cells/cm²). Cells were cultured in 96-well plates with 100 µl of *Medium I* as described. Values represent mean±SD of counting results from 2 wells each measured twice. To control for growth reducing effects caused by starvation, toxicity of metabolites and pH drop the endpoint cell density was verified and confirmed by completely independent replicates (subsequent experiment with new cryostock) with daily feeding with 200 µl medium (medium change every 24 h) contrary to feeding every other day (48 h) with 100 µl during the culturing for growth curve determination. To achieve a recommended optical confluence of 80 to 100 % at the time of transfection around 15.000 to 20.000 cells have to be seeded the day (20 ± 4 hrs) before in 100 µl of medium. Phenol red indicated that using 100 µl of medium (equal to the amount of transfection medium) no medium change is necessary within 48 hrs.

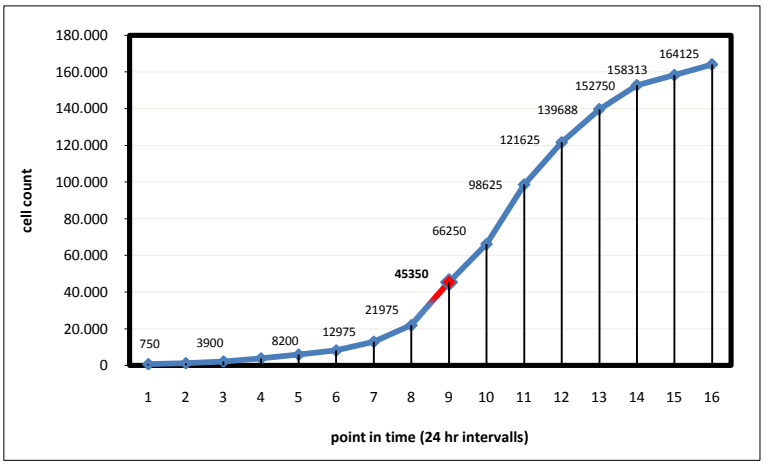


Fig. R2 HeLa growth curve for the 96-well format. The region where cell density approaches *optical confluence* is highlighted in red.

For the HeLa cell line (**Fig. R2**) optical confluence corresponds to about 35.000 to 40.000 cells per well (116.000 to 133.000 cells/cm²). Cells were cultured in 96-well plates with 100 µl of *Medium I* as described. Values represent mean±SD of counting results from 2 wells each measured twice. To examine the expected influences on growth caused by starvation, toxicity of metabolites and especially pH drop the plateau cell density was verified by completely independent replicates (subsequent experiment with cells from a new cryostock) with daily feeding with 200 µl medium (medium change every 24 h) contrary to feeding every other day (48 h) till point 8 and every day (24 h) beyond with 100 µl during the culturing for growth curve determination. The results are shown in **Fig. R3**. To achieve a recommended optical confluency of 80 to 100 % at the time of transfection around 15.000 to 20.000 cells have to be seeded the day (20±4 hrs) before in 100 µl of medium. In the region of interest (point 7 to 11) cell density is low enough to be properly supplied with 100 µl of standard medium for 24 hrs. Phenol red indicates that using 100 µl of medium (equal to the amount of transfection medium; though, transfection medium per well effectively contains less than 10% serum and less L-Glu) a medium change would be necessary after approximately 36 hrs. Although cells typically get lysed 48 hrs post transfection and viability seems to be stable over this time (7-AAD and trypan blue tested) a medium change 24 hrs post transfection or the addition of another 100 µl of complete medium may be advisable for maximal translation activity.

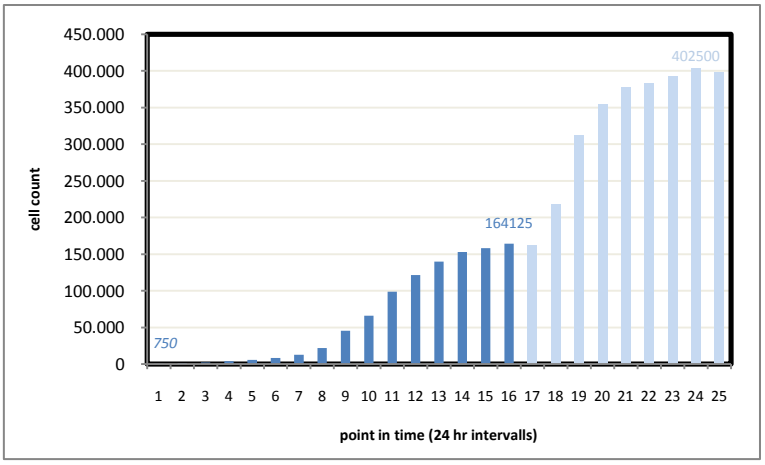


Fig. R3 HeLa growth curve for the 96-well format. The point of this figure: Keep the culture conditions constant, especially with HeLa cells.

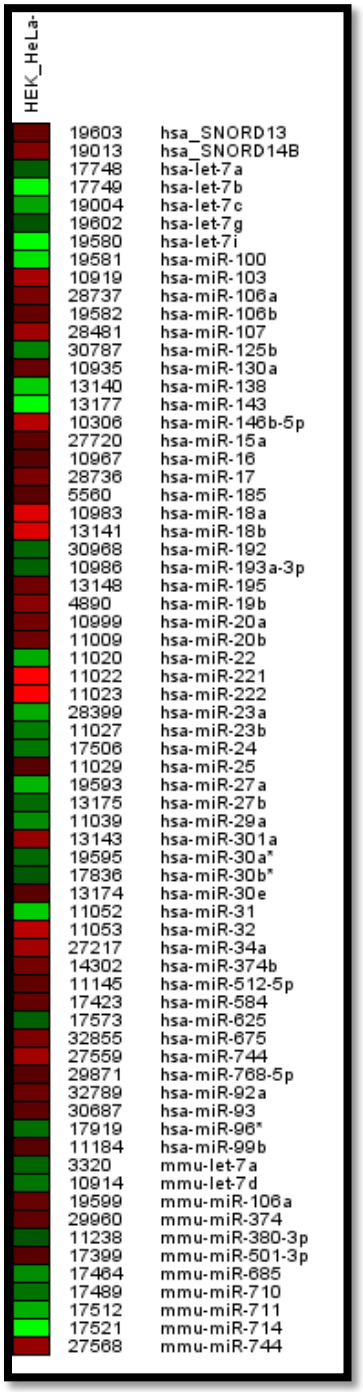
HeLa cells (as a cancer cell line) do not undergo real contact inhibition (stop of cell division due to mechanical signals), but divide until environmental factors like nutrient supply or pH-value (Ceccarini & Eagle 1971) force them to stabilize population density with periodical (e.g. daily) changes of medium (e.g. 100 µl medium until point 16; 200 µl beyond; **Fig. R3**) what explains the less sharp transition into plateau phase. In other words, reducing the feeding interval to less than 24 hrs or increasing the amount of medium (or portion of serum and especially L-Glu) would further increase the cell density and accelerate growth speed, respectively. Of course, cells do not grow as monolayer beyond point 10.

Tab. R1 Seeding recommendations. Shown are calculated values based on the number of seeded cells and the approximated average growth rate per hour at the respective cell density (derived from growth curves).

	*seeded cells (**HEK293**)*			*seeded cells (**HeLa**)*		
hours	**20.000**	**18.000**	**16.000**	**20.000**	**18.000**	**16.000**
16	37460	33714	29968	32094	28885	25675
17	38958	35062	31166	33057	29751	26446
18	40516	36465	32413	34049	30644	27239
19	42137	37923	33710	35070	31563	28056
20	43822	39440	35058	36122	32510	28898
21	45575	41018	36460	37206	33485	29765
22	47398	42659	37919	38322	34490	30658
23	49294	44365	39435	39472	35525	31577
24	51266	46139	41013	40656	36590	32525
25	53317	47985	42653	41876	37688	33500
26	55449	49904	44360	43132	38819	34505

Strictly speaking, the seeding recommendations (**Tab. R1**) only apply when healthy, regularly subcultured, well growing cells are used and supplied with fresh FBS and L-Glu at seeding time (prepare aliquots, store at -20°C). Of course, these numbers do not consider settling time and cycle perturbation due to transfer induced stress. Nevertheless, cell growth typically shows only little variance, and in general these recommendations work very well in reaching 90 to 100 % optical confluency after the indicated number of hours. Indeed the main variance causing influence on cell density at transfection time might be inaccuracies in cell counting prior to seeding.

4.2 miRNA expression profiling



Everything is relative. Also in the cases where the enrichment or depletion of particular miRNAs is the adjusting wheel in the experimental setup this saying holds true. Expression at a moderate level makes it easier both to repress and enhance a miRNA activity. Adding 1 amole of a functional molecule to 1 amole of its kind already working in the cellular environment may be expected to show a stronger effect on the adjusted system states than adding it to 10 amoles typically accomplishing their mission, likewise a certain amount of an inhibitor may draw a clearer picture when it encounters rather less than more representatives of its target molecule population. Additionally, it may be favored to have a rather low background abundance of possibly interfering family members. In cases where expression levels are sufficiently different (certainly given in all 68 listed here) the knowledge of cell type specific expression also permits target testing without the need to transfect miRNA mimics or inhibitors – what may not be the method of choice in general as endogenous miRNA levels can change on many reasons during the assay time, but it can serve as cheaper pre-test or additional evidence.

For these reasons, a miRNA expression profiling on the two model cell lines currently available in-house for luciferase assays (HEK293 vs. HeLa) has been conducted (**Fig. R4**).

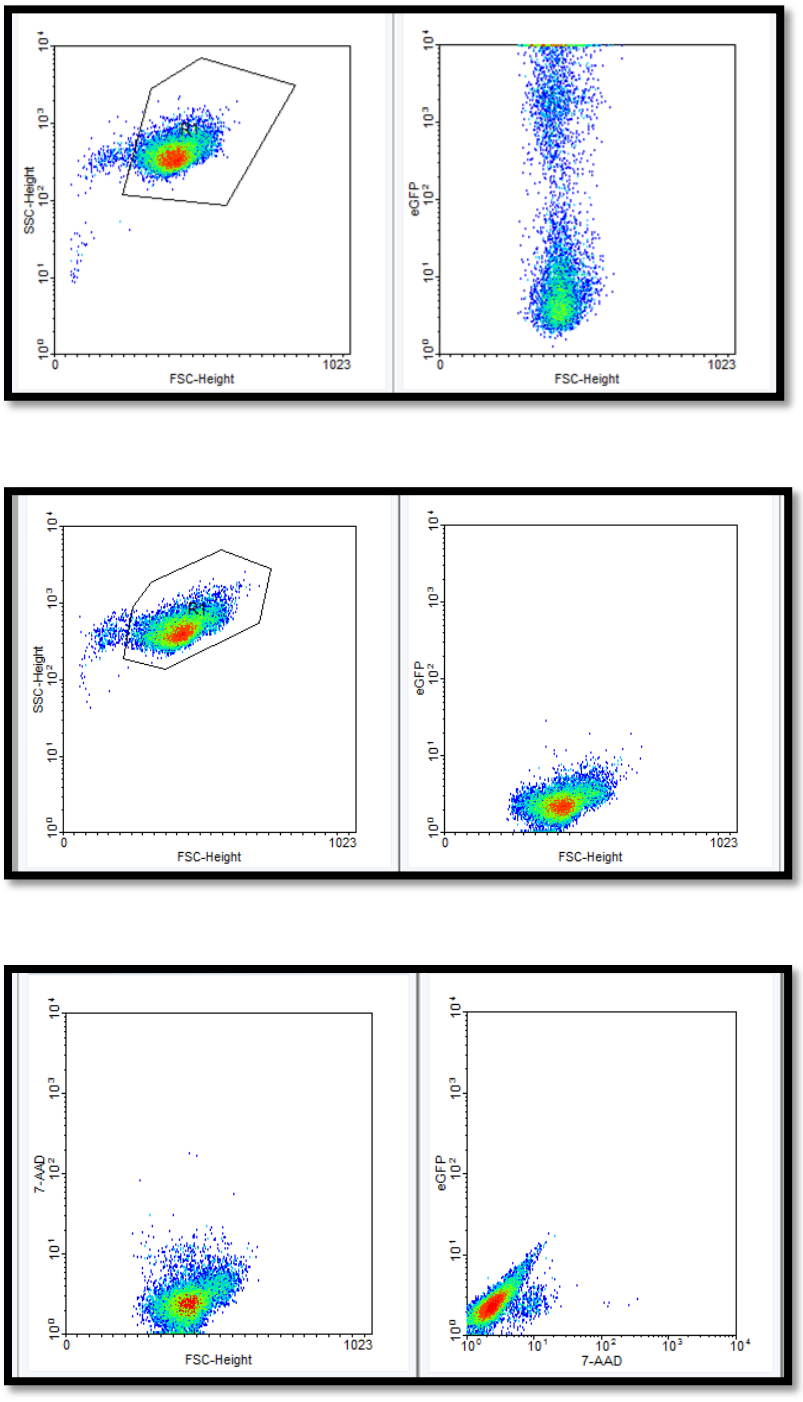
Fig.R4 miRNA profiling – HEK293 vs. HeLa as reference.

68 miRNAs show differential expression ($\log_2$ ratio above absolute 1) between HEK293 and HeLa cells. RNA samples were taken from nearly optical confluent proliferating cells in log growth phase using *TRIzol*.

For example, the studied miR-27b (see below) has a $\log_2$ ratio of -1.28 (data not shown). In other words, the level of miR-27b in HEK293 cells is about 60% lower than in HeLa cells.

4.3 Flow cytometry results

4.3.1 Representative scatterplots and intensity distributions



Gate R1 excludes all registered events with an FSC-Height value (representative for particle size) or SSC-Height (representative for particle granularity) outside the typical distribution for the particle (cell) species. Therefore, in the eGFP/FSC plot no events in the lower FSC-Height channels appear (they would represent cell debris caused by sample preparation). The high event density around the intensity level of 10^1 represents $eGFP^-$ and dim shining $eGFP^+$ cells, whereas all events above 10^2 clearly show $eGFP^+$ cells.

For comparison, the pictures in row two show untransfected cells (controls).

7-AAD positive cells are clearly to identify, especially in the eGFP/7-AAD plot on the right they are prettily set off against the auto-fluorescence distribution.

Fig. R5 Representative scatterplots showing primary gating for exclusion of cell debris and eGFP/7-AAD positive examples.

Though, determination of eGFP$^+$ and 7-AAD$^+$ cells can be done in scatter plots, intensity distributions (**Fig. R6**) have been used for discrimination of respective positive or negative cells (scatter plots have been used for controls).

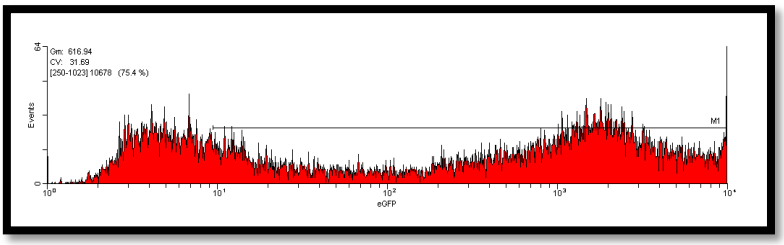


Fig. R6 eGFP$^+$ screening – representative example of an intensity distribution. The identifiable normally distributed events between 10^0 and 10^1 on the logarithmic eGFP intensity scale (x-coordinate) represent autofluoresence of the non-transfected cell population and partially dim shining eGFP$^+$ cells (disturbing the bell curve on the right). The rather conservative setting of the marker M1 (determining the proportions) in this example is representative for all analyses conducted, thus true eGFP$^+$ percentage might be slightly higher than stated (for the reason why this does not really matter see discussion on eGFP expression). Nevertheless, presented FACS analysis results reveal the relative best transfection conditions. Determination of 7-AAD$^+$ cells works equally.

4.3.2 Substitution of siRNA/miRNA by plasmid DNA

Chemically synthesized siRNAs and miRNA mimics are rather expensive, thus for initial FACS studies and determination of optimal conditions with respect to cell viability and *liposome-nucleic acid* ratio (indeed, fine-tuning has been conducted using luciferase and luminometer measurements, and different amounts of sRNA of course) they have been substituted for *pBluescript II KS-* (BSC; Stratagene, Cat#212208) a plasmid devoid of eukaryotic promoters. Though, it is not generally advisable to do this as it is known that (in dependence of the particular lipid used) sRNA and pDNA complex formation may be quite different (Spagnou et al. 2004), after consulting the technical support of both manufacturers this approach was used. Possible effects of substitution on *lipid-nucleic acid* complex formation and hence transfection efficiency and viability have been tested for both lipid-based transfection reagents, with the result that differences have been minimal and not statistically significant (5% α-level). Though, in theory even differences in the composition of particular plasmids may have an influence on efficiency and toxicity of a transfection, in practice it is often possible to maintain once determined optimal transfection conditions by simply keeping the ratio between the total amount of nucleic acids (both DNA and RNA) [µg] to lipid solution [µl] constant – this only holds true when the format (e.g. 96-well) is maintained. However, it is known that while large DNA molecules do not change physicochemical characteristics or morphological and structural features of lipoplexes they negatively affect gene transfer capacity (and thus protein amounts) – in other words, gene transfer is most effective with small plasmids (< 20 kb), and small molecules in general (providing higher copy numbers) (Kreiss et al. 1999), thus plasmids differing substantially in size may be expected to show different signal intensities in luminescence measurements.

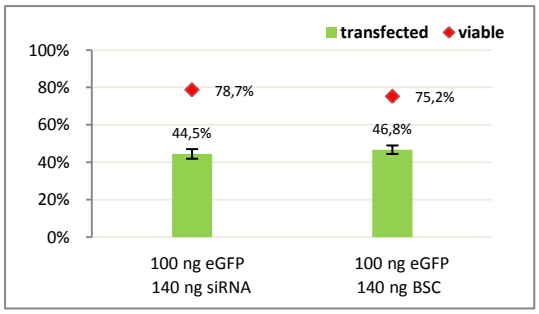


Fig. R7 Effect of sRNA substitution by plasmid DNA. Shown are means±SD of three biological replicates. *Setup*: DFD/NUC ratio: 1.7; MP/NUC ratio: 3 (data not shown); post-transfection incubation time: 48 hrs; cell line: HEK293; t-test: $p_{transfected} \sim 0.5$, $p_{viable} \sim 0.14$

4.3.3 Validation of viability values

To check whether the obtained viability values determined by 7-AAD staining are reliable, samples were split and a parallel *trypan blue* staining was conducted. In addition, the influence of different amounts of eGFP-plasmid transfected per well were examined to assess toxicity effects of protein expression.

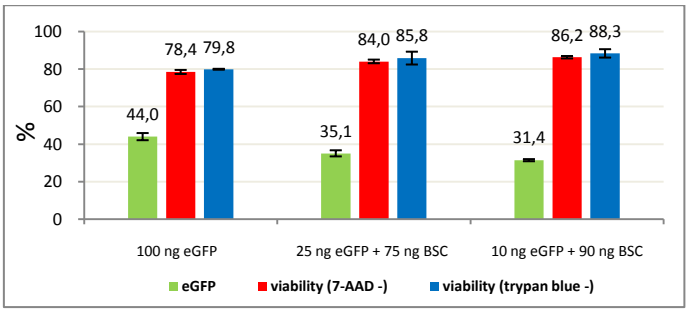


Fig. R8 Validation of 7-AAD results by *trypan blue* staining – representative for HEK293 and DFD. *Setup:* cell line: HEK293; 18.000 cells seeded 20 hrs prior; DFD/NUC ratio: 1.7; post transfection incubation time: 48 hrs

Some GFP expressing cells contracted and rounded up (also with HeLa cells), as reported by Lui et al., thus in subsequent experiments only 25 ng of eGFP have been used, where the rest was filled with inactive *pBluescript II* (BSC) (H S Liu et al. 1999). The seemingly lower transfection efficiency with less eGFP is caused by a higher proportion of dim shining $eGFP^+$ cells that fell under the threshold for being identified as positive, related to the fact that the plasmid copy number per transfected cell heavily determines the expression rate (Tseng et al. 1997, Kreiss et al. 1999). Apparently, proportions of viable cells indicated by 7-AAD and *trypan blue* agree well.

4.3.4 Determination of optimal *lipid-nucleic acid* ratios and *lipoplex amount-cell density* relation

The morphology of *lipoplexes* (*lipid-nucleic acid* complexes) depends on the cationic lipid and the lipid composition (e.g. neutral lipids) of the particular transfection reagent and on the proportion of nucleic acid to lipids. Each cell line has specific compatibility to certain lipoplex morphology which cannot be predicted, thus the optimal lipid-DNA/RNA ratio has to be determined experimentally to achieve both high transfection efficiency and viability of cells – being crucial to reliability and reproducibility of results. Generally it is assumed that lipoplexes have to feature a positive net charge to allow attracting electrostatic interactions with cell membranes. Indeed this is the most important point in the establishment of lipid-based transfections. A second important factor is the amount of lipoplexes applied per cell, where too less clearly lowers transfection efficiency and too much causes intolerable toxicity (see **Fig. R19**) (Biontex).

The red line in the following figures (**Fig. R9** to **Fig. R14**) indicates the lowest proportion of viable cells seen as acceptable by both manufacturers (Dharmacon, Biontex) for reliable experimental outcomes. Shown are mean proportions of transfected ($eGFP^+$) and viable ($7\text{-}AAD^-$) cells in the gated cell population (gate G1 see **Fig. R5**) over lipoplex amount and *lipid-nucleic acid* ratio (x-axis description: first row, nucleic acid amount; second row, lipid amount; third row, *lipid-nucleic acid* ratio). Error bars indicate standard deviations of 2 technical replicates.

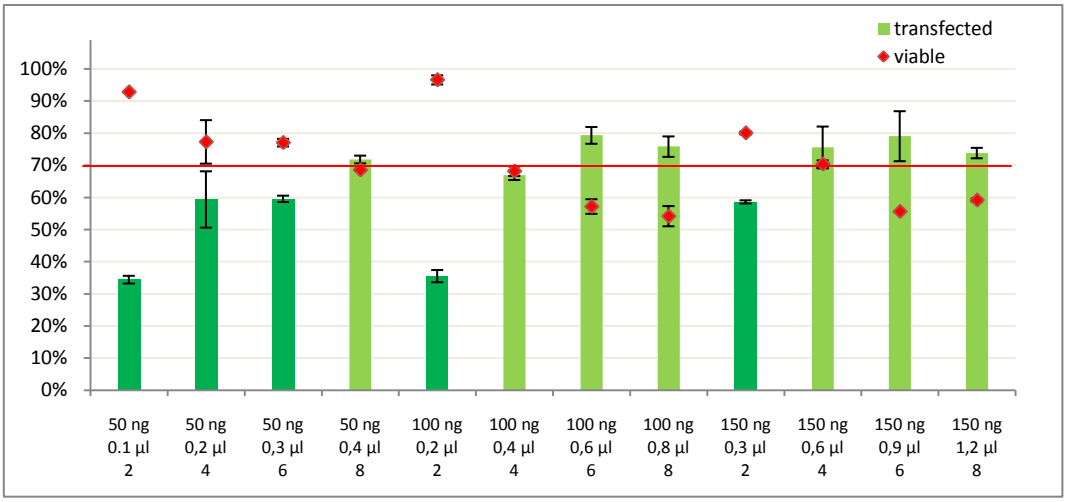


Fig. R9 DharmaFECT Duo (DFD) transfection – 10.000 cells seeded.

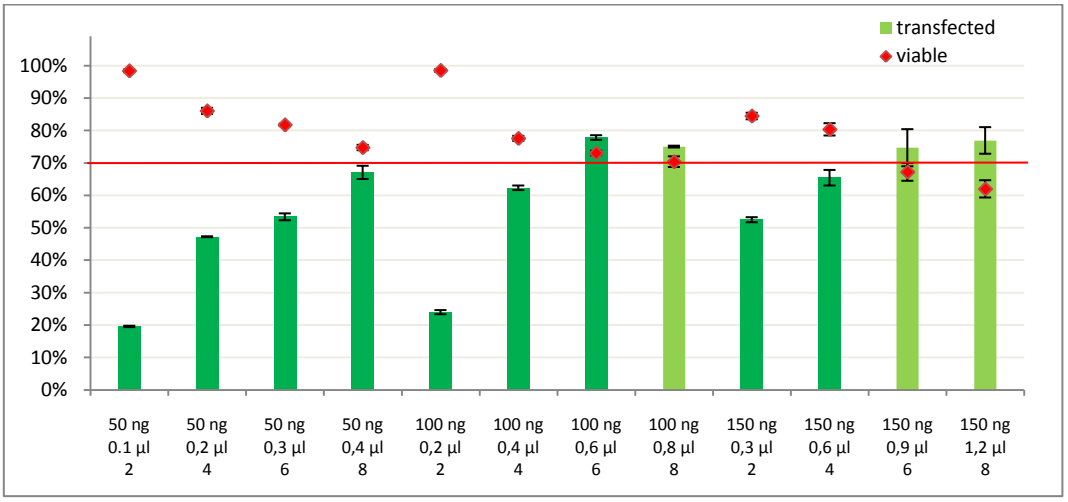


Fig. R10 DharmaFECT Duo (DFD) transfection – 20.000 cells seeded.

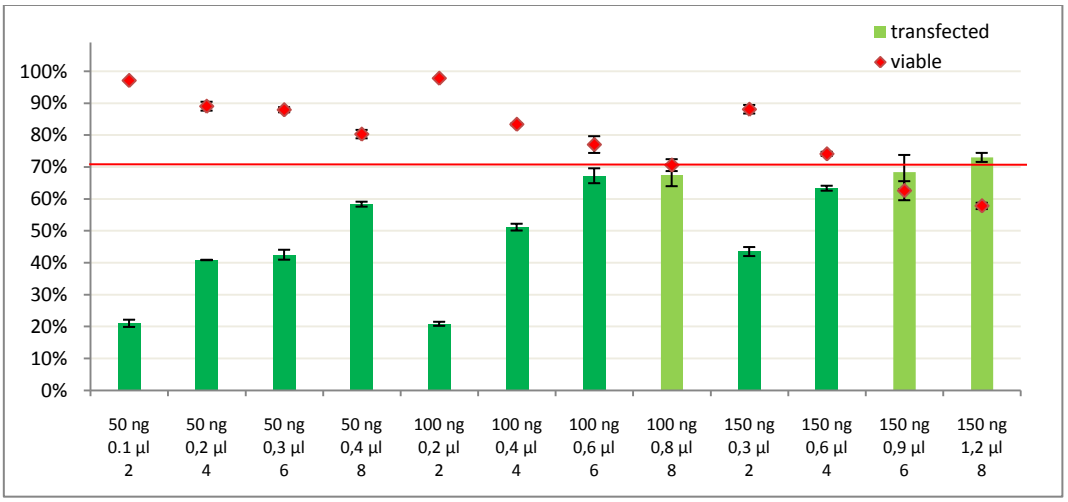


Fig. R11 DharmaFECT Duo (DFD) transfection – 30.000 cells seeded.

Taking Dharmacon's recommendation as basis up to 100 ng of plasmid(s) and up to 50 nM of miRNA mimic (summing up to ~ 165 ng nucleic acid) shall be transfected in a standard 96-well assay, where both values may be titrated down once the optimal ratio(s) have been found. Thus, a range of 50 to 150 ng of DNA was transfected to check for amount effects on viability and transfection efficiency. Lipoplexes were formed with different ratios of lipid to DNA [µl/µg] to find out at which ratios better transfer/transgene expression occurs. Furthermore, the same transfection media were spread on 3 different cell densities (with respect to growth curves) to see influences of this parameter – cells were seeded 20 hrs prior to transfection. FACS analysis was done 48 hrs post transfection.

In DFD transfections (**Fig. R9** to **Fig. R11;** *dark green:* possible conditions) a higher proportion of lipid in the lipoplexes clearly shows increased toxicity in the examined ratio range of 2 to 8 with all amounts of lipoplexes on all cell densities, where a ratio of 8 shows no more improvement as compared to a ratio of 6 – indeed, there is an indication of already declining efficiency. Increasing cell density at seed respectively transfection tendentially decreases cytotoxicity, but lowers proportions of transfected cells what becomes substantially with 30.000 cells at seed. These results are in accordance with expectations. In all cases – even with 30.000 cells – viability drops below the threshold level when more than 0.6 µl DFD are used (what is the maximum recommended by the manufacturer for the 96-well format). Thus, there is an indication that at high lipoplex amounts rather low ratios of 2 to 4 are favorable, where at lower amounts also ratios up to 6 are applicable with respect to viability. Considering both transfection efficiency and viability a cell seed density of ~ 20.000 may be optimal for most lipoplex amounts.

In MP transfections (**Fig. R12** to **Fig. R14**; dark green: possible conditions) there is no obvious trend of increasing toxicity with increasing lipid proportions (apparently Biontex' *toxicity optimization module* approach works). The effect of decreasing lipoplex toxicity and decreasing transfection efficiency with increasing cell density is more pronounced with MP as compared to DFD, thus correct cell seeding is more important. The marked drop in ratio 6 and 8 conditions with increasing amounts of cells cannot be explained, as the same transfection media were applied to all cell densities – but the results for 20.000 and 30.000 cells may be trustworthy as optimal ratios of 2 to 4 for HEK293 cells are also described in the literature. Declining efficiency with increasing ratios after the optimum is not untypical for lipid-based transfections, but would rather be expected to be accompanied by worse viability values. Though not obvious by checking membrane integrity cellular processes may be disturbed as inspection by light microscopy show cellular stress with MP amounts > 0.7 µl (see **Fig. A8**). Very clearly ratios of 2 to 4 over a broad range of lipoplex amount on ~ 20.000 cells at seed may be optimal.

As against HEK293 cells for HeLa cells manufacturer and different user recommendations were available for both reagents this initial test was omitted in its full for this cell line – however, small-scale tests (with 10.000, 20.000 and 30.000 cells) were conducted, indicating that also for Hela cells a density of ~ 20.000 cells at seed (20 hrs) and a broader range of lipoplex amount are performing well at ratios of 2 to 4 – with both reagents (higher ratios were not tested as in subsequent experiments 250 ng of DNA/RNA were to be transfected what already demands 1 µl of lipid per 96-well – although, MP shows less toxicity in HeLa cells as compared with HEK293 cells more than 1 µl would likely be adverse as under light microscopic inspection abnormal phenotypes are already visible with 1µl MP and with respect to HEK293 results and manufacturer's statement high amounts of DFD might also be adverse).

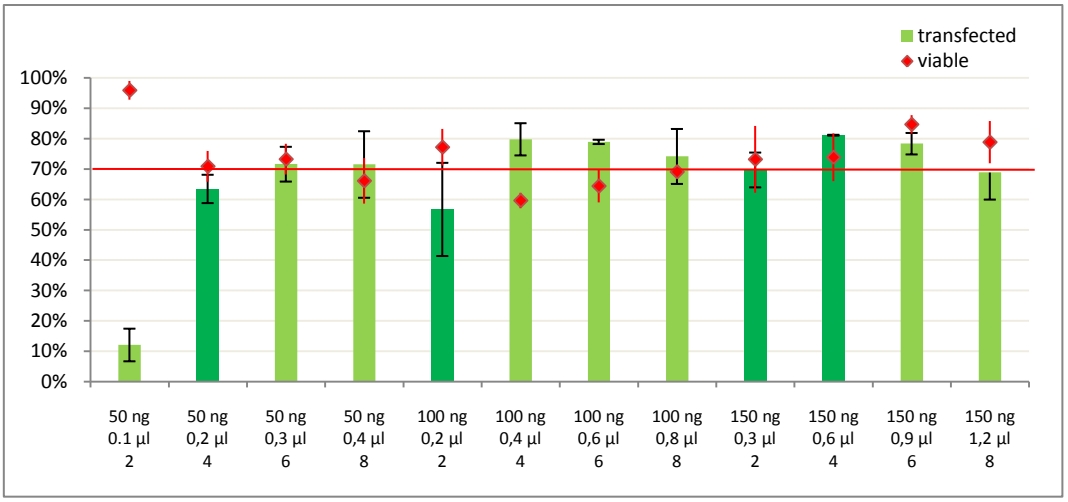


Fig. R12 METAFECTENE Pro (MP) transfection – 10.000 cells seeded.

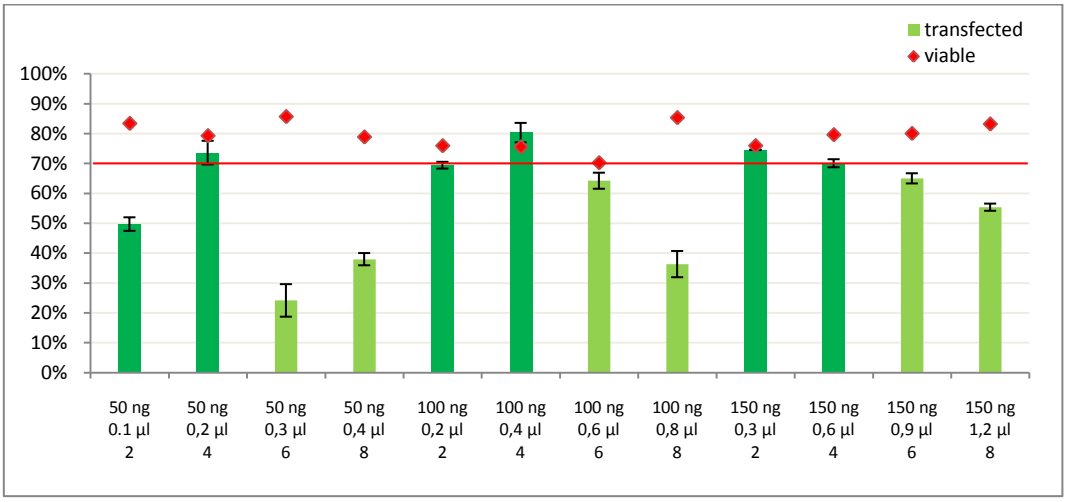


Fig. R13 METAFECTENE Pro (MP) transfection – 20.000 cells seeded.

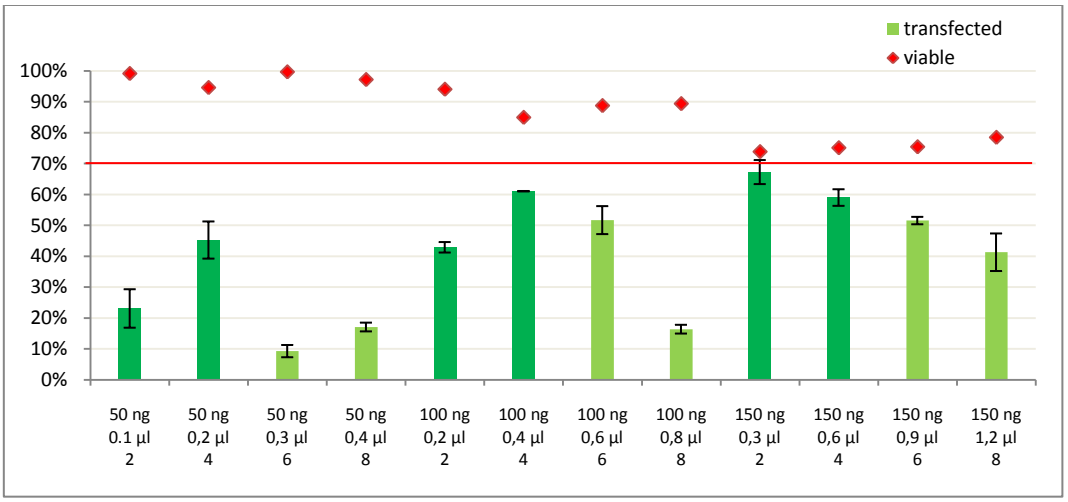


Fig. R14 METAFECTENE Pro (MP) transfection – 30.000 cells seeded.

Having set the cell density at seed to ~ 20.000 (seeded ~ 20 hrs prior transfection; in accordance with estimations from growth curves and general recommendations) for both cell lines (see **Fig. A10** and **Fig A11**, Appendix) and reagents another ratio check with a fixed DNA amount of 250 ng was conducted, representing the transfection of 100 ng plasmid(s) and ~ 100 nM sRNA in total (e.g. mimic and inhibitor or two mimics).

In all cases (**Fig. R15** to **Fig. R18**) 18.000 cells were seeded 20 hrs before transfection of 250 ng DNA (25 ng eGFP + 225 ng *pBluescript II*, approximating 100 ng of *psiCHECK-2* and 100 nM (10 pmol) of siRNA/miRNA) with MP and DFD, and measured 48 hrs post transfection. Again, manufacturers' recommendation to use only conditions where cell viability is higher than 70% is indicated in red.

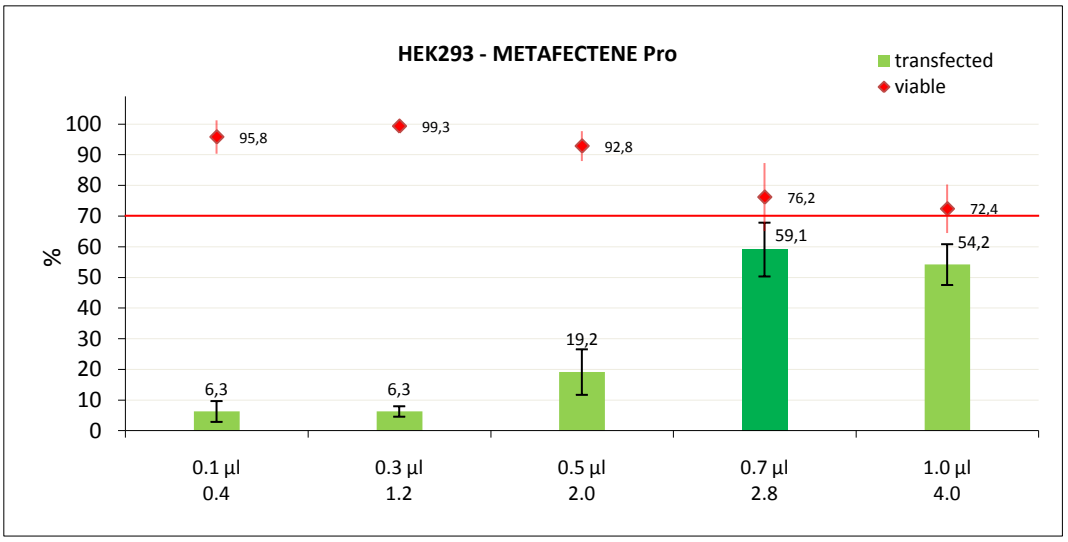


Fig. R15 Transfection efficiency and viability of HEK293 cells transfected with METAFECTENE Pro.

For MP and HEK293 cells the optimal ratio of reagent to nucleic acids is about 3 (**Fig. R15**), where the achieved maximal efficiency of ~ 60% for high lipoplex amounts has also been identified for this cell line by several independent (quantitative) studies (Biontex; Application notes). The drawbacks compared to DharmaFECT Duo (**Fig. R16**) are the higher variability of results and the apparently higher cytotoxicity of MP-based lipoplexes at this lipoplex amount. Indeed, MP transfections have been the only ones of all conducted that exhibited clearly visible cell stress at reagent amounts of > 0.7 µl under optical microscope inspection (see **Fig. A8**, Appendix). But, in general the transfection system should be capable to transfer ~ 250 ng of nucleic acids (corresponding to 100 ng of luciferase vector and 100 nM of siRNA/miRNA), thus sticking to a ratio of ~ 3 demands 0.7 µl MP per well.

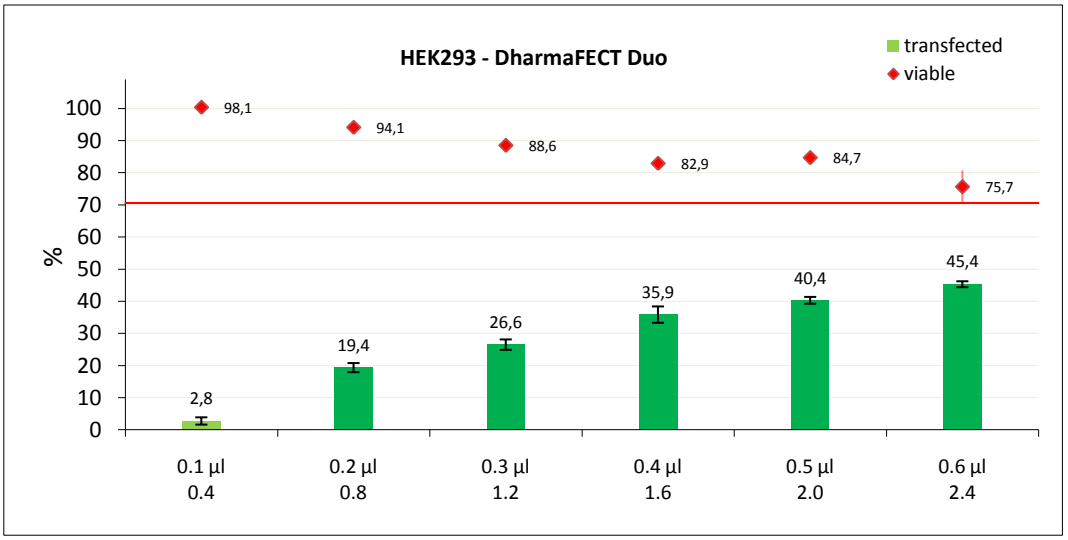


Fig. R16 Transfection efficiency and viability of HEK293 cells transfected with DharmaFECT Duo.

With HEK293 cells DFD shows a broad range of applicable lipid/NUC ratios from 0.8 to 2.4 (beyond viability might decline to a level below 70%; as 0.6 µl per 96-well are the upper limit recommended by the manufacturer no higher ratio has been tested in this setup, but **Fig. R10** indicates the trend) and transfection results are exceptional stable (**Fig. R16**).

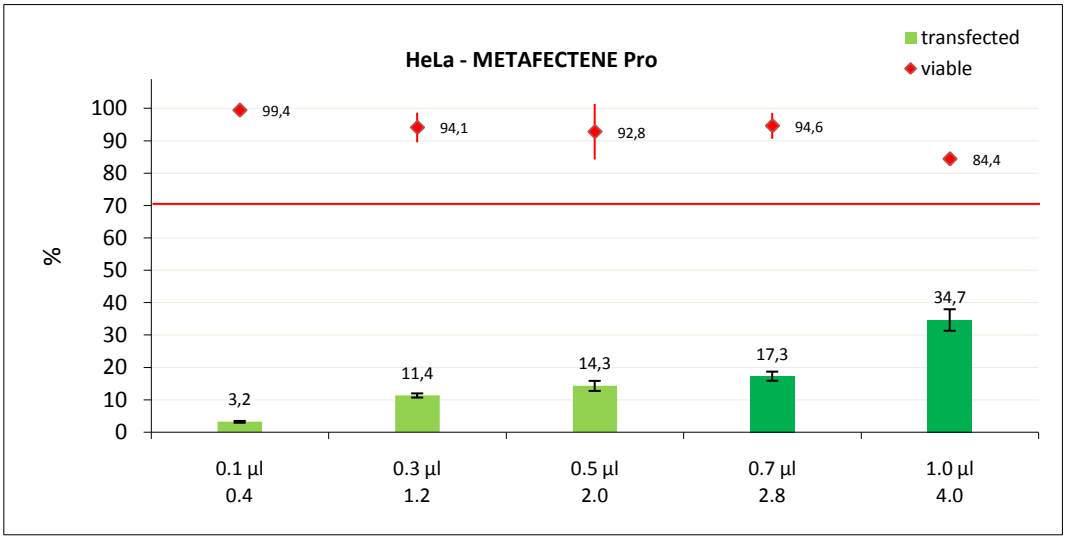


Fig. R17 Transfection efficiency and viability of HEK293 cells transfected with METAFECTENE Pro.

In HeLa cells the optimal ratio of MP-reagent to nucleic acids is about 4 (**Fig. R17**), where the achieved maximal efficiency of ~ 35% lies in the range of 30 to 50% that has been identified for this cell line by independent quantitative studies (Biontex; Application notes). The variability of MP results is comparable to DFD results, and cytotoxicity seems to be lower according to membrane integrity checks. Though, 1 µl MP caused visible cell stress (see **Fig. A8**).

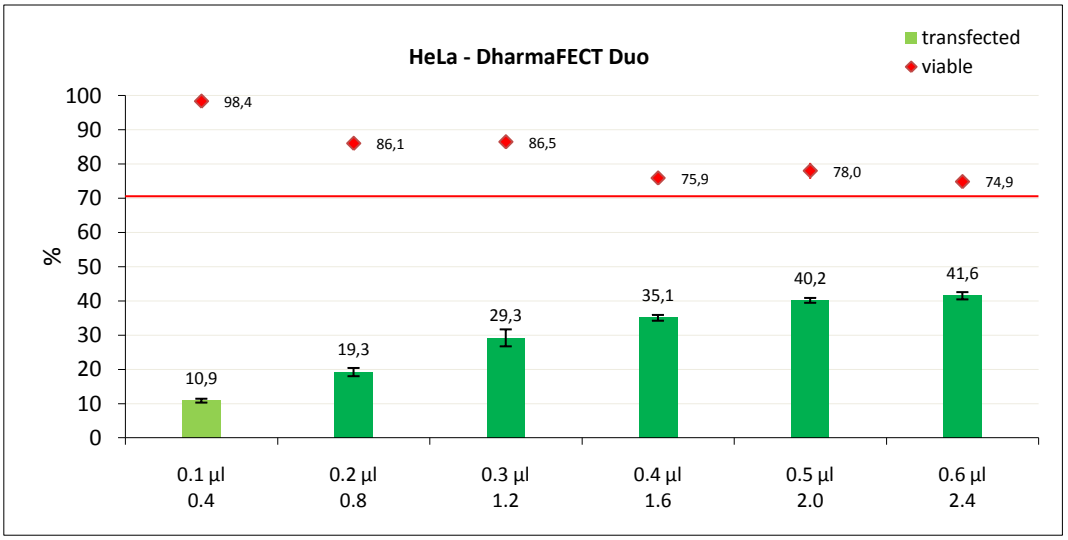


Fig. R18 Transfection efficiency and viability of HeLa cells transfected with DharmaFECT Duo.

Also in HeLa cells DFD shows a broad range of applicable lipid/NUC ratios, again, basically from 0.8 to 2.4 with very stable transfection results (**Fig. R18**). Small-scale experiments indicated that with both cell lines MP transfections perform slightly better in the presence of serum (as stated by the manufacturer) and DFD transfections perform slightly better *without* serum in the first 6 hrs after transfection (~ 3 to 5 % increase in transfection efficiency determined by eGFP) but with an concomitant loss of stability (increase of standard deviations by a factor of ~2; too high a price). Thus, all transfections (MP and DFD) were conducted with 10% serum present. Indeed, based on FACS analysis results both transfection reagents could be used for the assay, but there are indications that the comparatively high *lipid-nucleic acid* ratios MP demands may become a problem when higher amounts of nucleic acids shall be transfected (e.g. co-transfections of several miRNAs).

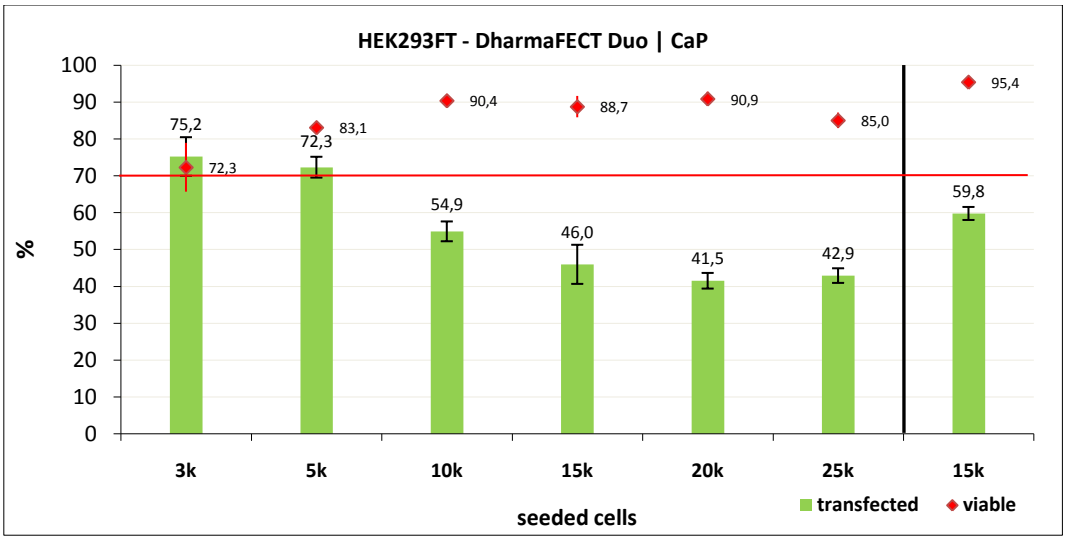


Fig. R19 Representative result of a complex amount-cell density relation – for comparison a CaP transfection result on the right. *Setup:* cell line:HEK293FT; seeded 18 hrs before transfection; transfected with 250 ng plasmid (25 ng eGFP, 225 ng BSC) and 0,3 µl DFD; CaP according to description in methods

4.4 Luminometer results

4.4.1 *Lipid-nucleic acid* ratio check

In all experiments 100 ng of *psiCHECK-2* and 100 nM of siRNA (*siGENOME* Non-Targeting siRNA Pool #2) have been transfected with the indicated amounts of the respective transfection reagent – the siRNA pool consists of 4 different siRNAs, where one targets the firefly luciferase variant contained in the vector. The following figures (**Fig. R20** to **Fig. R23**) represent transfection efficiency tests analogously to the aforementioned FACS analyses with 240 ng DNA. RLU values for *RL* indicate the extent of transgene uptake and expression, values for *FL* additionally depict the successful transfer of siRNA (see below).

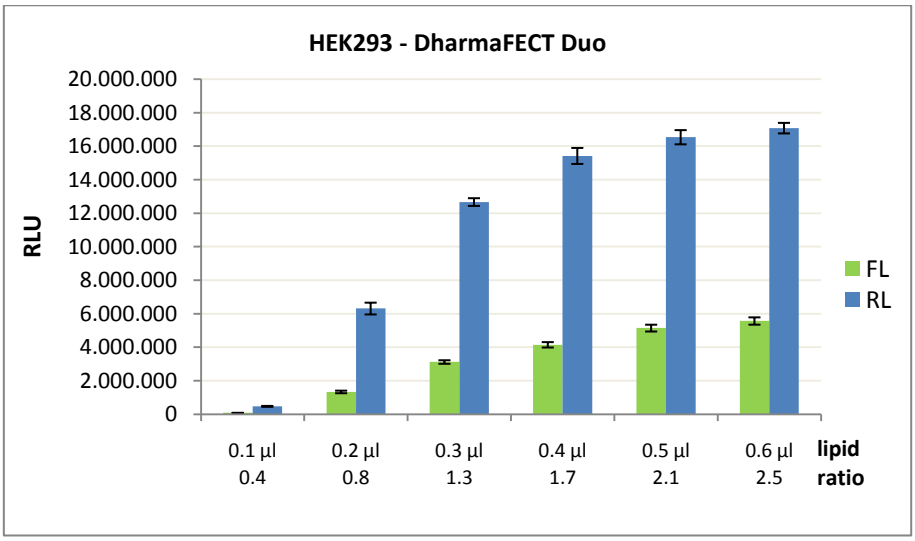


Fig. R20 DFD/NUC ratio – HEK293. *Used:* 5 µl sample + 25 µl reagent

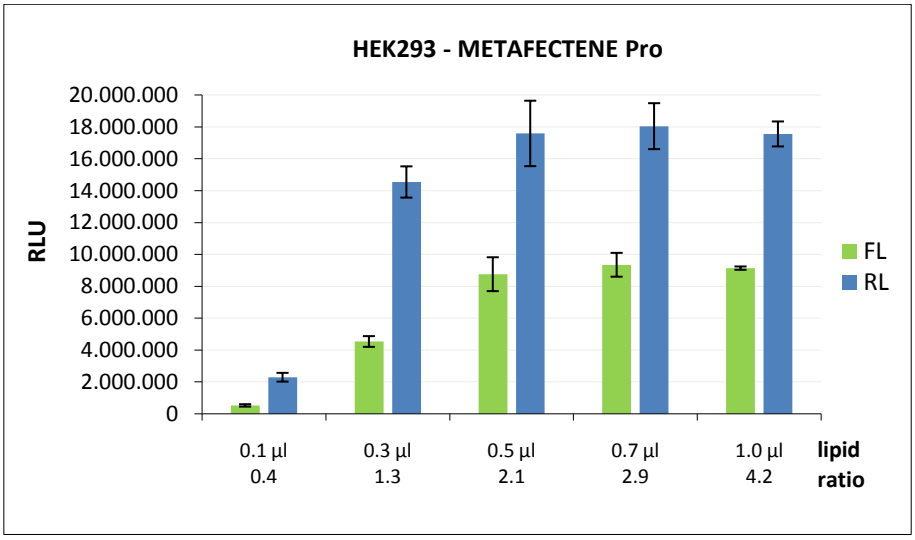


Fig. R21 MP/NUC ratio – HEK293. *Used:* 10 µl sample + 50 µl substrate

As already indicated by FACS results a DFD/NUC ratio range from 0.8 to 2.5 is applicable and data exhibit very low variance (**Fig. R20**). Also in accordance with FACS results (for 240 ng DNA) variance of MP transfections is higher than with DFD (**Fig. R21**). Strikingly, RLU values in dependence of ratio do not go together with eGFP$^+$ values over ratio (compare **Fig. R15**) – especially the efficiency jump at ~3 to 4 is not present. Apparently ratios of 2 to 3 work best, the slight decrease in the ratio 4.2 bar may correspond to the decrease in FACS results and indicate toxicity effects (as mentioned 1 µl MP caused already visible stress on both cell lines, see **Fig. A8**).

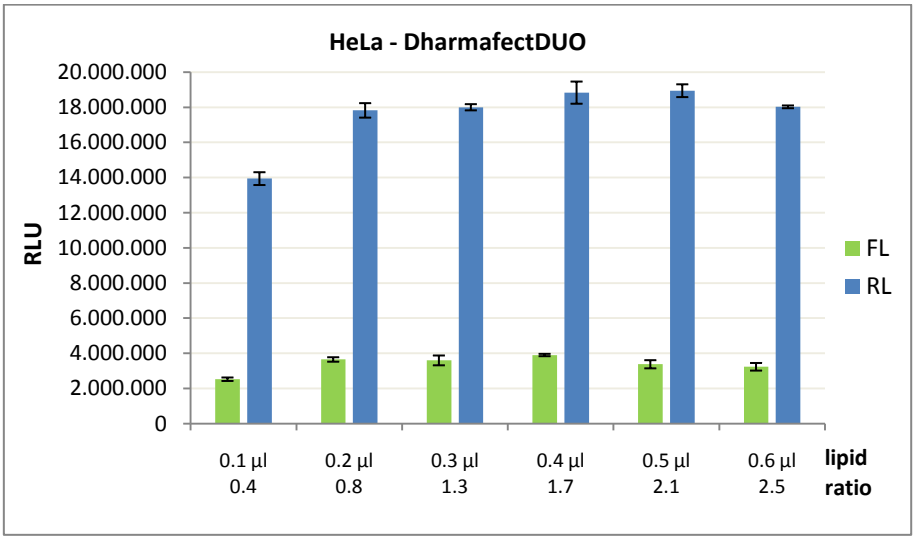


Fig.R22 DFD/NUC ratio – HeLa. *Used:* 5 µl sample + 25 µl reagent

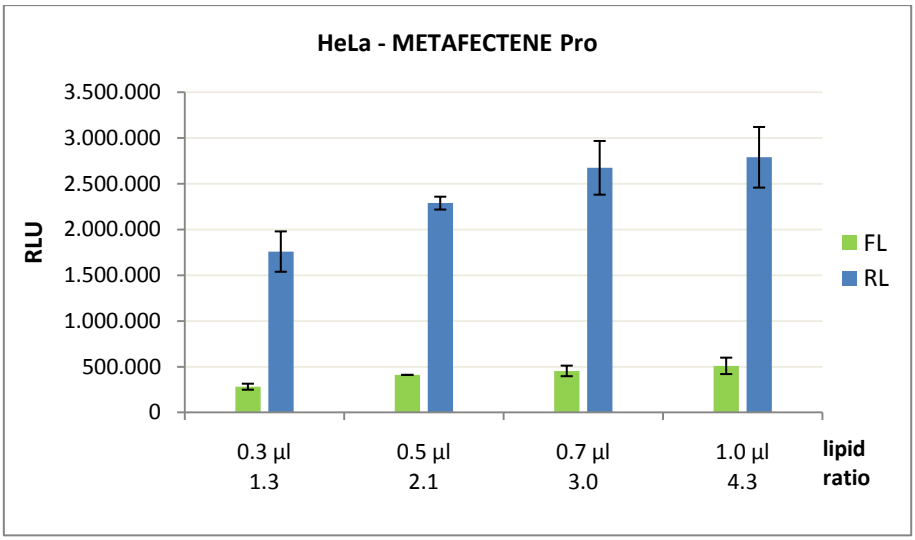


Fig. R23 MP/NUC ratio – HeLa. *Used:* 10 µl sample + 50 µl substrate

For HeLa cells treated with DFD the smooth increase of positively transfected cells with increasing ratio seen in FACS analysis carries over to luminescence results (**Fig. R22**). As with HEK293 cells a

broad range of DFD/NUC ratios is applicable (0.4 to 2.5) and variance of technical replicates is very low.

With MP-transfected HeLa cells, as already seen in HEK293 transfections, the eGFP expression (again especially the jump at a ratio of 4; compare **Fig. R17**) did not reproduce in the luminescence measurement (**Fig. R23**). This may be due to vector specific effects in combination with MP – probably related to eGFP characteristics (see discussion on GFP expression). But, as MP transfections with subsequent luminescence measurements have been conducted several times these results can be seen as reliable. Thus, ratios of 2 to 4 may be applicable.

For both cell lines and both transfection reagents the luciferase ratios are stable over different *lipid-nucleic* acid ratios (CV values < 5%, no tendency; data not shown), thus in the examined range different composition of lipoplexes with respect to lipid and nucleic acid proportions apparently does not influence the knockdown strength substantially – for evaluation of knockdowns in each experiment (**Fig. R20 – Fig. R23**) a separate vector only (without siRNA) transfection was conducted (not shown above; MP/DNA ~2, DFD/DNA ~ 1.7).

Although in principle still both reagents are applicable it is obvious that transgene transfer/expression is weaker in MP transfections, especially with HeLa cells (be aware of the fact that with MP transfected samples two times the amount of cell lysate was measured for luminescence). Moreover, substantially greater variance of measured RLUs in technical replicates do not militate for MP.

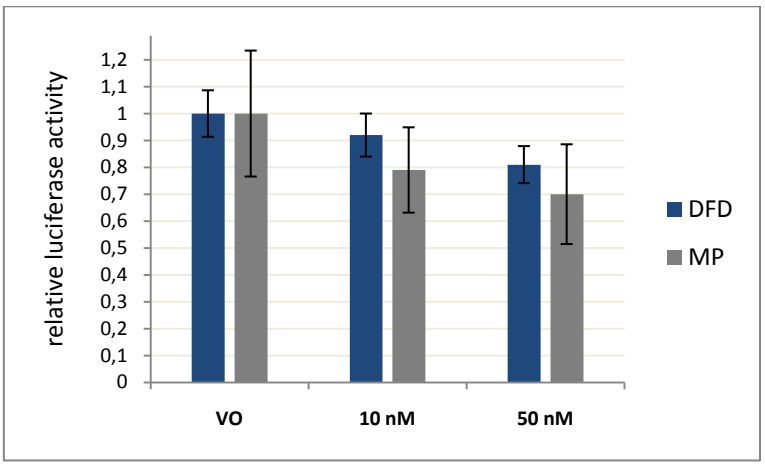


Fig. R24 Knockdown comparison – DFD vs. MP. Error bars represent standard deviations of 4 samples where always two were transfected with the same transfection mix (2 x 2 technical replicates). *Setup:* 18.000 HEK293 cells seed 20 hrs prior transfection; DFD/NUC ~ 1.8; MP/NUC ~2; siRNA: *siGENOME* Non-Targeting siRNA Pool #2, vector: *psiCHECK-2*

Also in the normalized (*RL*/*FL*) ratios data from MP-transfected cells show much more variance as compared with DFD results (propagation of uncertainty applied) (**Fig. R24**). Supposably MP is very sensitive to little handling mistakes (like slightly different incubation times or losses of liposomes due to vessel surface contact). However, in the viewed sRNA molarity range until 50 nM siRNA the knockdown in MP-transfected samples is about twice as strong as in DFD-transfected ones for 10 nM

and still 50% higher with 50 nM. Though, with respect to statistical tests smaller knockdowns in connection with much smaller variances are clearly favorable against the opposite.

4.4.2 Post-transfection incubation time – luciferase expression and knockdown

Both luciferase enzymes and miRNA mimics/inhibitors have life-times as any other cellular regulator or actor. To determine the degradation behavior in the standard incubation time frame for protein analyses (72 hrs) changes in knockdown strength and RLU values were traced. The utilized interaction pair RB1 and miR-26a was previously validated as functional by *Michael Karbiener* (IGB) who also kindly provided the RB1 construct.

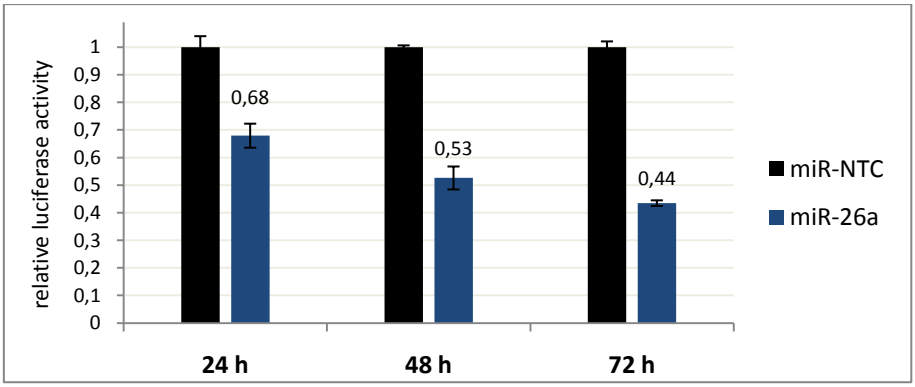


Fig. R25 Timing knockdown – HEK293. *Setup:* 80 ng plasmid RB1 construct; 50 nM miR-26a or miR-NTC; 18.000 cells seeded 20 hrs prior.

Although the standard incubation time after transfection was chosen with 48 hrs, **Fig. R25** indicates that at least in some cases (construct/miRNA pairs) knockdown strength may further increase with prolonged incubation.

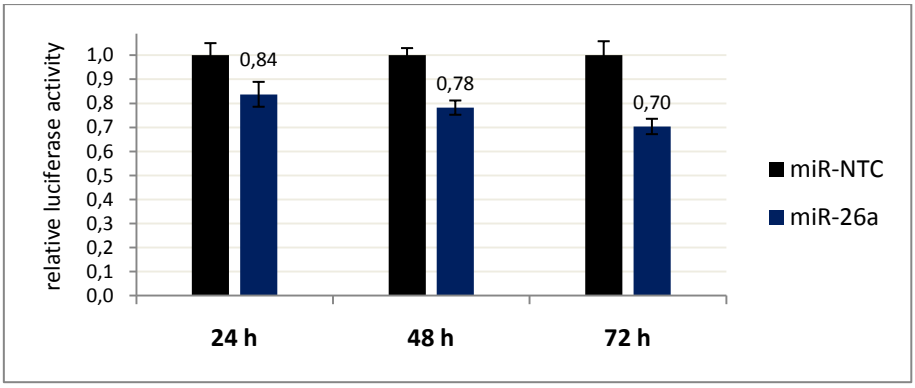


Fig. R26 Timing knockdown – HeLa. *Setup:* 80 ng plasmid RB1 construct; 50 nM miR-26a or miR-NTC; 18.000 cells seeded 20 hrs prior.

Compared with the HEK293 result weaker knockdowns appear in HeLa cells (probably due to endogenous miR-26a) but the same trend of increasing knockdown over 72 hrs is present (**Fig. R26**). Thus, also in HeLa cells 72 hrs incubation may be favorable for this combination.

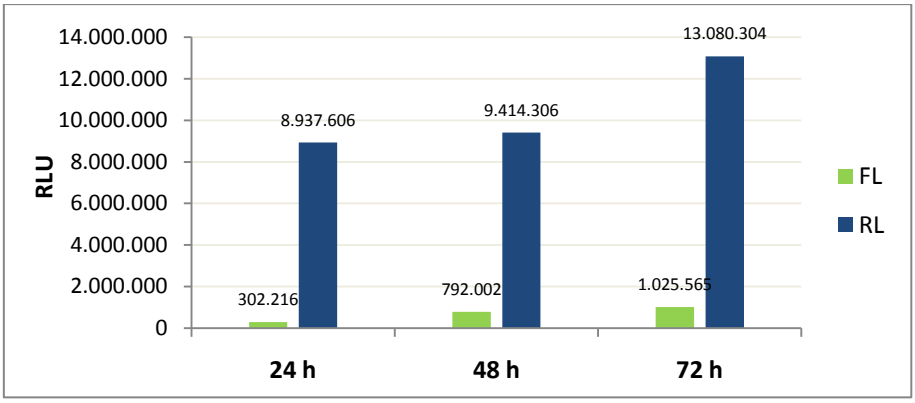


Fig. R27 Timing RLU values – HEK293. *Setup:* 80 ng plasmid RB1 construct; 50 nM miR-NTC; 18.000 cells seeded 20 hrs prior.

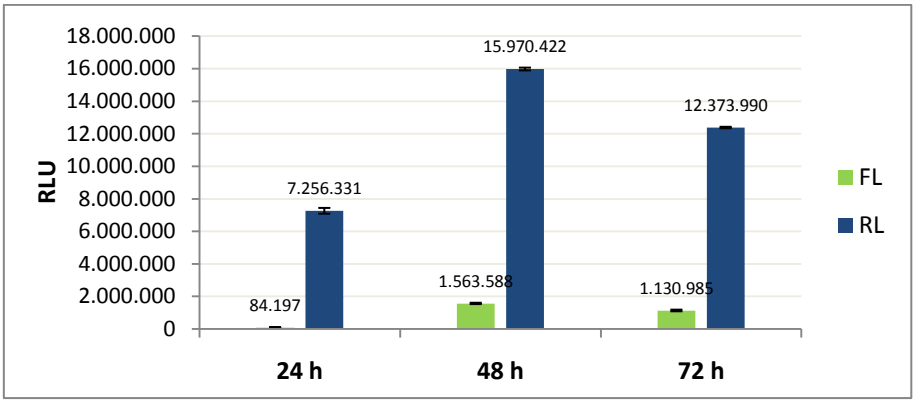


Fig. R28 Timing RLU values – HeLa. *Setup:* 80 ng plasmid RB1 construct; 50 nM miR-NTC; 18.000 cells seeded 20 hrs prior.

For the same construct the protein amount was traced over 72 hrs by luminescence measurements (luciferase activity) with the result that concentrations of both luciferases stably increase over time in HEK293 cells (**Fig. R27**). In contrast, the maximal luciferase concentrations in HeLa cells are reached after 48 hrs. The corresponding samples shown in **Fig. R25** to **Fig. R28** were treated with the same transfection media and handled equally until measurement, thus variations are cell line specific.

4.4.3 Effect of sample and substrate amount reduction

Besides the chemically synthesized small RNAs, the substrates of the two luciferases are the most expensive materials used in this assay. Thus, a reduction from the manufacturer's recommendation of 100 µl substrate (for 20 µl luciferase containing cell lysate) down to 50 µl or even 25 µl is reasonable from an economic point of view. As in virtually all cases 0.5 to 5 µl of cell lysate produce an acceptable (or even too high, see section **4.4.4 linearity concerns**) signal level a corresponding downscaling of the used substrate amount is no problem with respect to keeping the luciferin concentration in excess. Though the amount of luciferase in a sample can vary widely, in all experiments conducted so far a substrate solution [µl] to sample amount [µl] ratio of 5 (as recommended by manufacturer) assured substrate being in excess. For peace of conscience one may measure the supposably strongest sample (with the highest expected luciferase concentration), e.g. some negative control in overexpression experiments or the sample treated with maximal inhibitor concentration in silencing experiments, in different substrate reagent to sample amount ratios to check on which point there arises substrate limitation and determine the minimum amount of necessary reagent, respectively (**Fig. R29**).

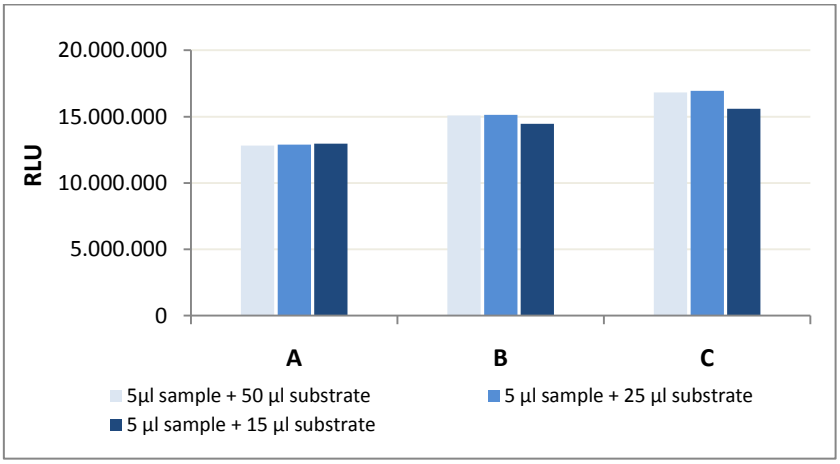


Fig. R29 Effect of reagent amount reduction due to luciferin limitation. Samples A, B, and C contain increasing concentrations of *Renilla* luciferase molecules. Each sample was measured with 3 different amounts of substrate reagent.

Obviously in the cases B and C substrate availability limits the enzymatic reaction rate when only 15 µl of substrate solution are added to 5 µl of sample. In case the tested sample A would be the strongest to be expected in a subsequent measurement session, all samples could be measured with only 15 µl reagent per well. Some care has to be taken of final buffer composition. Following manufacturer's recommendation the ratio of PLB (sample) to first injected *LARII* (the *FL* reagent; buffer and substrate) is 1:5, and accordingly 1:10 to second added *Stop&Glo* (the *RL* reagent; buffer and substrate). Several experiments indicate that initial ratios of 1:3 to 1:10 (as in **Fig. R29**) do not compromise buffer conditions and enzymatic reaction, respectively. Although RLU values decrease with relatively increasing amounts of PLB, the RLU ratios remain constant in that range, but it should be avoided to go below 1:3 as slight (and not necessarily uniform) increases in RLU ratios are likely – so, do not dilute a strong 1 µl sample in 19 µl of PLB and add it to 25 µl of reagent, if there are

concerns about appropriate sample-substrate mixture, but dilute either in 4 to 9 and respect pipetting instruction below or dilute in 9 to 14 µl and use 50 µl of reagent.

Everything has its price, so does economizing. Although 25 µl of fluids with a viscosity comparable to that of water do cover the area of a well (96-well Assay Plate; Costar, Cat#3912) entirely, the *LARII* buffer does *not* due to higher viscosity and different contact angles (may change when a different substrate kit or plates with different surface characteristics or made of another material are being used). That is why care must be taken when pipetting a low amount sample (< 15 µl cell lysate) into a well if only 25 µl of substrate or less is going to be used. Due to insufficient mixing of luciferase and luciferin containing buffers variation in the results may be the consequence (**Tab. R2**). In the case of 25 µl the error introduced for this reason is expected to be below 15 %, but it can easily be avoided by pipetting routinely approximately into the center of a well (even with 3 µl sample). It is not generally advisable to scale down further as results may become unreliable – for instance, using 3µl sample and 15 µl substrate may end up in a measurement error of up to 50 % if no care is taken about appropriate pipetting.

	relative errors in RL/FL ratios [%]		
	hline	***vline***	***center***
RE_{min}	0.05	4.04	0.08
$RE_{25\%}$	0.97	4.54	0.23
$RE_{50\%}$	1.32	5.08	0.37
$RE_{75\%}$	1.59	7.35	0.42
RE_{max}	2.23	13.04	0.46

Tab. R2 Relative error (RE) in dependence of sample placement. Low amounts of sample (< 15 µl) in combination with low amounts of substrate (< 30 µl) can cause errors-prone results due to incomplete mixing caused by the physical properties of the substrate buffers and the setting of the injection system of the luminometer. To overcome this problem one has to pipette into the center of a well (error < 0.5 %) or at least on a point in the horizontal axis (*hline*) crossing the center (error < 2.5 %). 5 different samples (3 µl; negligible amount) have been measured each 2 times in 5 positions (center and the four intersections of the well wall and a thought plate seams aligned cross through the center) using 25 µl of each luciferase substrate. The REs of a certain position (like the center) were calculated in relation to the mean of all values for this position and sample, whereas the (clearly systematic) REs within a well (*hline, vline*) are related to the center value of the particular well, because of the fact that the RE distribution (with RE_{max} < 0.5 %) of repeated measurements of center placed samples is virtually the same using 25 µl or 50 µl (or even 100 µl) of substrate solution.

4.4.4 Luminometer issues – linearity concerns

Normal background signal levels caused by electronic circuits, dark current (release of electrons in the PMT's photocathode not as consequence of the photoelectric effect but due to thermal movement of atoms), autofluorescence of plates (may increase when exposed to sunlight for extended periods) and reagent buffer (especially caused by the *Coelenterazine* in the *Stop&Glow* reagent) are below 200 RLU. In typical assays the lowest signals are well above 100.000 RLU and even in less successful experiments with bad transfection or weak translation of the construct, RLUs hardly fall below a value of 10.000. Basically, one can measure down nearly to the background level (3 standard deviations above background), which should be determined separately in each experiment with a lysate of untreated cells.

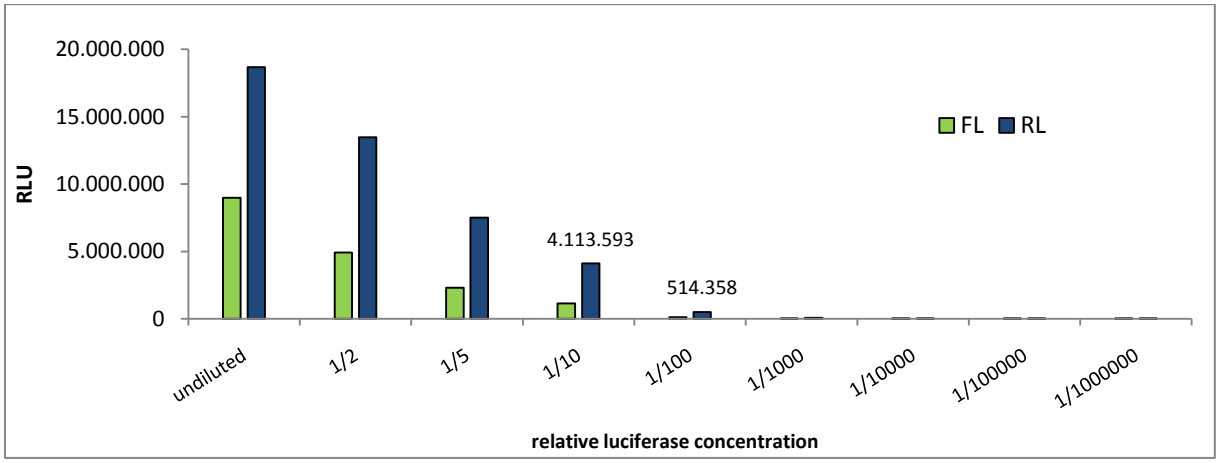


Fig. R30 Serial dilution - checking dynamic range and luminometer linearity. Serial dilution was conducted with *passive lysis buffer* (PLB; Promega, Cat#E1941), where 1mg/ml gelatin has been added to prevent the loss of the luciferase enzymes from solution due to absorption at great dilutions.

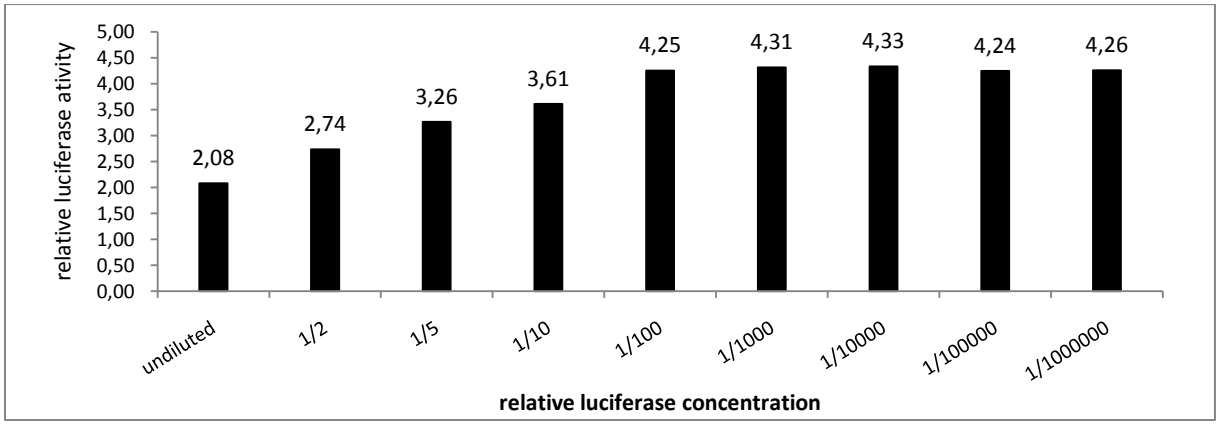


Fig. R31 Effect of PMT nonlinearity on RLU ratios.

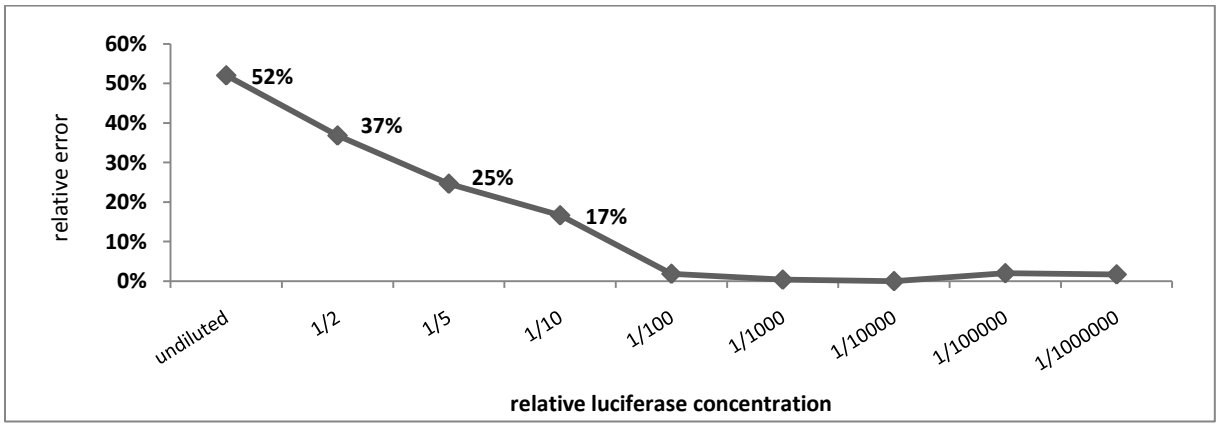


Fig. R32 Relative error in *RL*/*FL* ratios as function of relative luciferase concentration and photon flux, respectively.

20 MIO RLUs are the highest possible measured value the device (*ORION II*, Berthold) displays, any photon flux beyond will give the output **4294967295** (2^{32}), indicating saturation of the PMT. Serial dilution was conducted to check the dynamic range and the linearity of the luminometer (**Fig. R30**). The CV of *RL*/*FL* ratios in the linear region (below 1 MIO RLU, here 1/100 and stronger diluted) is despite of supposable imprecision in the dilution procedure below 1 % (**Fig. R31**). Outside the region of linear relation between *RL* and *FL* signals the ratio decreases mainly due to erroneous photon counting (measurement beyond resolution limit of the PMT and indeed coincidently incoming photons). As non-linearity of the device has a stronger effect on the (higher) signal of the higher concentrated luciferase (usually *RL* when using *psiCHECK-2*) a lowering of the RLU ratio is the consequence. The 1/10 dilution corresponds to about 4 MIO RLUs truly measured in the higher expressed luciferase (*RL*) and shows a deviation in the ratio of approximately 17 % what should be seen as maximal tolerable with respect to standard deviations, significance tests and influences on calculated knockdowns of fusion-mRNAs (**Fig. R32**). Thus, samples with higher raw values have to be diluted (ideally below 2 MIO RLU) to obtain correct results.

The obvious non-linearity is caused by the fact that the PMT is operated in single photon counting mode. As the incoming of single photons is a statistical process and photons possess a certain pulse duration the probability of two or more photons arriving at the same time (truly or as seen by the PMT which cannot resolve the exact arrival time below a certain threshold and counts as a single event/photon) or superimposing increases with photon flux (light intensity). With respect to pulse resolution the delay time of the PMT and associated electronics is determining – for the actual luminometer this time is around 10 to 15 ns. This means that with 1 MIO cps (counts per second) delay time losses are in the region of 1 to 1.5 % (at 10 MIO ~ 20%, at 20 MIO ~ 40%) (personal communication; Berthold, technical support).

4.4.5 Shifted patterns

In biological replicates it may happen that there is a shift in relative luciferase activity patterns, the definitive reason for that is yet unknown. Obviously it is not due to substrate degradation or injection errors (e.g. air bubbles in the injection system) as in these cases the different luciferase concentrations would be hit with different strength. But a likely reason is a too low temperature (not room temperature) of either sample or reagents (see **5.2.1.3**), lowering enzyme activity. As this shift introduces additional variance it has potentially adverse influence on significance tests.

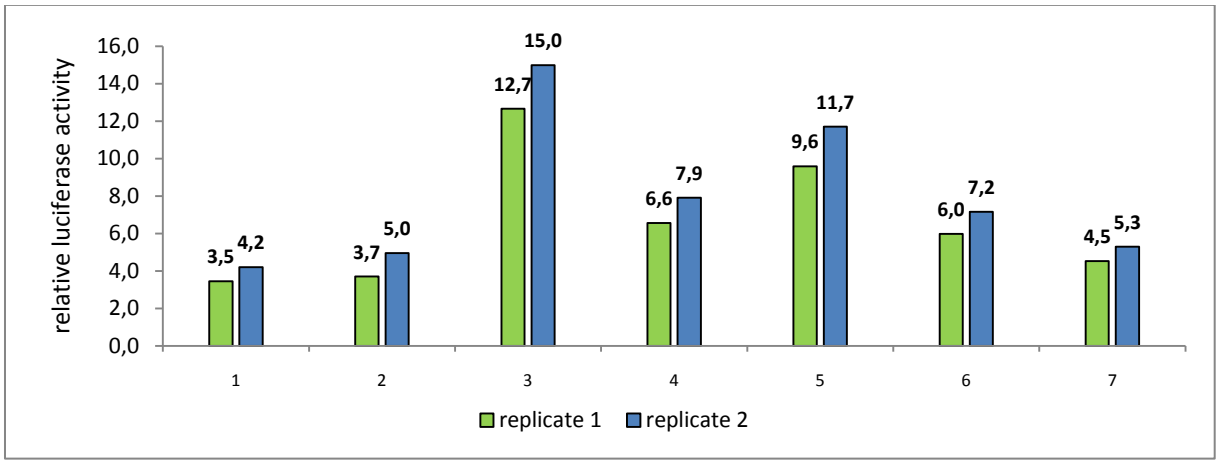


Fig. R33 Shift in relative luciferase activities of two biological replicates. In this example ratios in the second experiment are about 20% higher in all samples.

Occasionally it may happen that in biological replicates the pattern of effects is preserved but has experienced some shift usually caused by constantly (over all samples or a subgroup) decreased levels of reference luciferase (*FL*). Although repeated measurement may solve the problem (if enough sample amount is still left), there is actually no need for that as data can be corrected. In order to create prettier yet correct results some scaling may be done. The principle of this mathematical cosmetic is a plain multiplication with a constant factor derived from all experiments. This factor may be the average of all factors between related ratios (mean shift) or the factor between a single related pair (or factors between n-tuples in case of n replicates). Both, mean and standard deviation simply have to be multiplied or divided by this/these factor(s) in relation to a freely selectable reference experiment. Alternatively, each single biological replicate can be normalized to a selected sample (usually some negative control) giving values relative to 1 and final statistic values may be calculated from these results. In any case error propagation of uncertainty must be used when calculating standard deviations.

4.4.5.1 Effect of correction

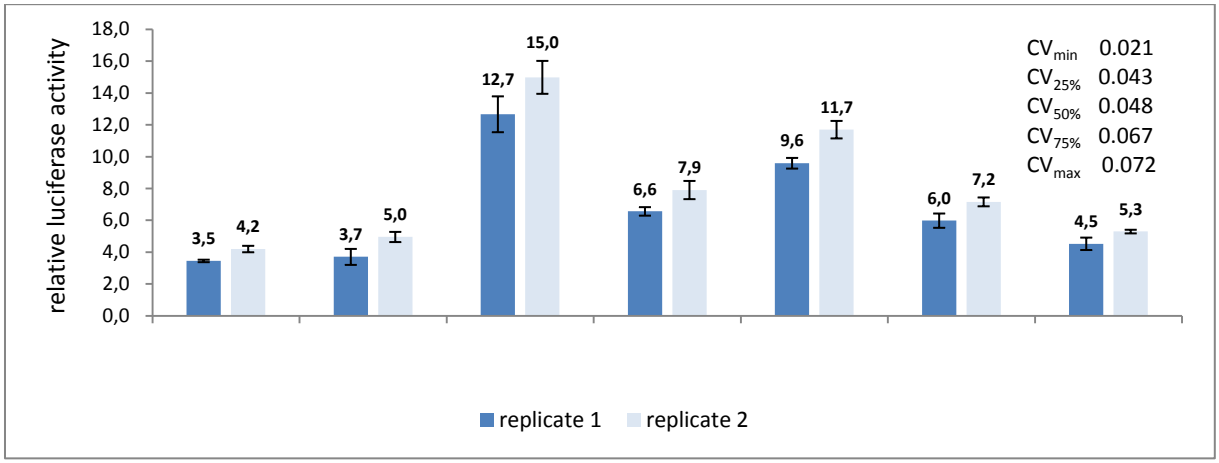


Fig. R34 Shifted effect patterns. Mean and standard deviations of 3 technical replicates per sample are shown for 2 biological replicates. The CV distribution (all values < 10 %) indicates proper execution of the experiment.

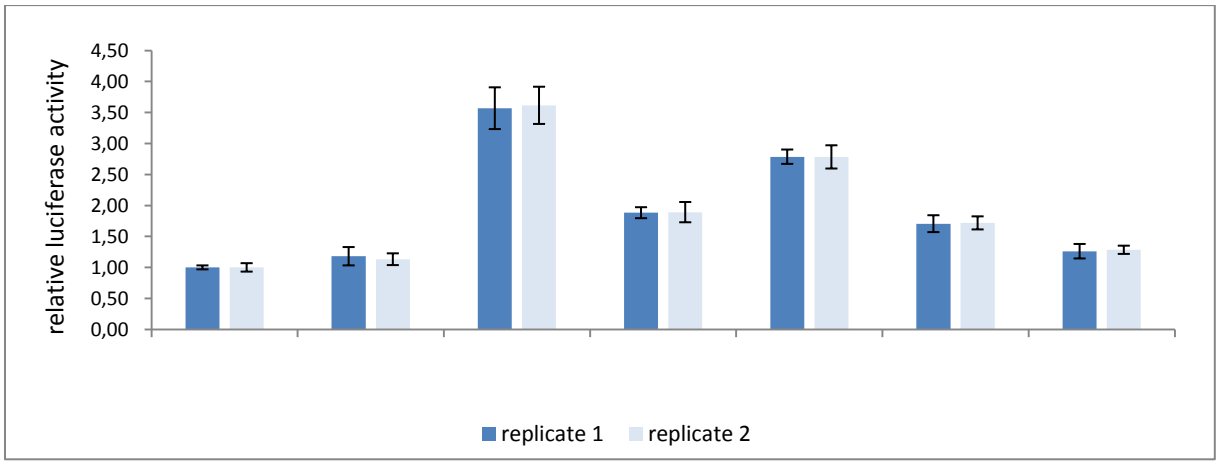


Fig. R35 Normalized effect patterns. Normalization reveals an exceptional well conformity of the replicated results with corresponding effects on error bars after joining the data (**Fig. R36** and **Fig. R37**).

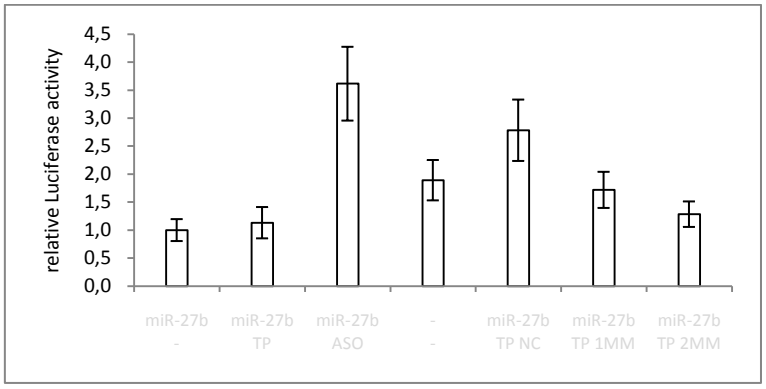


Fig. R36 Effect of result scaling. Standard deviation before correction.

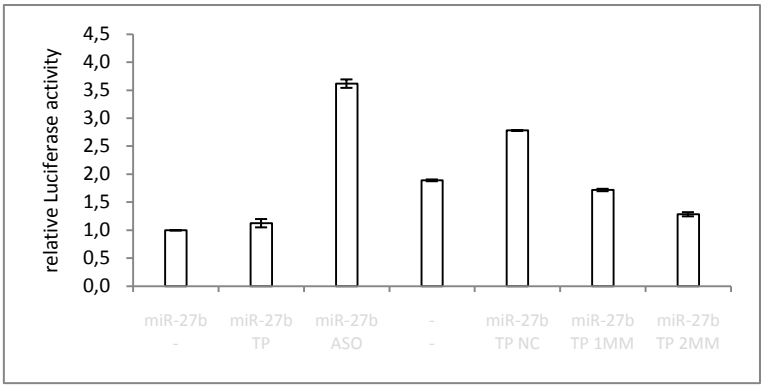


Fig. R37 Effect of result scaling. Standard deviation after correction.

The reduction in standard deviation (and due to the small sample sizes typically observed effectively the reduction of the standard error) is very relevant with respect to statistical tests for significance of observed effects – especially with low mean differences.

4.4.6 Stability and reliability – technical replicates and linearity

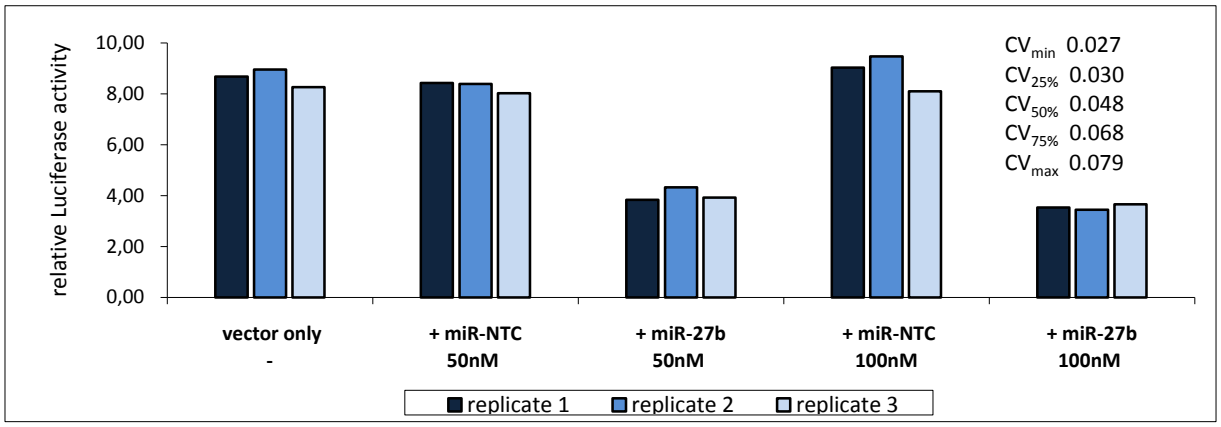


Fig. R38 Typical pattern of technical replicates.

Technical replicates shall be understood as separate wells of a 96-well plate treated with the same transfection mix at the same time and sharing the same environmental conditions during treatment and incubation, but being independent in all other potentially variance causing parameters like exact cell density and viability, contamination, pipetting, or final luminometer sample preparation and measurement. CV values of technical replicates, representing the cumulative variation (and error) introduced, are usually (and should be) below 10% - depending on the handling skills of the experimenter. Larger deviations are mostly caused by effects of PMT nonlinearity, for instance, the third bar in the *+ miR-NTC 100nM* group (responsible for CV_{max}) is more out of line because the higher raw value is in the region of 5 MIO RLUs compared to rather 3 MIO RLUs in the other two replicates – sample dilution and repeated measurement would be necessary to eliminate this variance (**Fig. R38**).

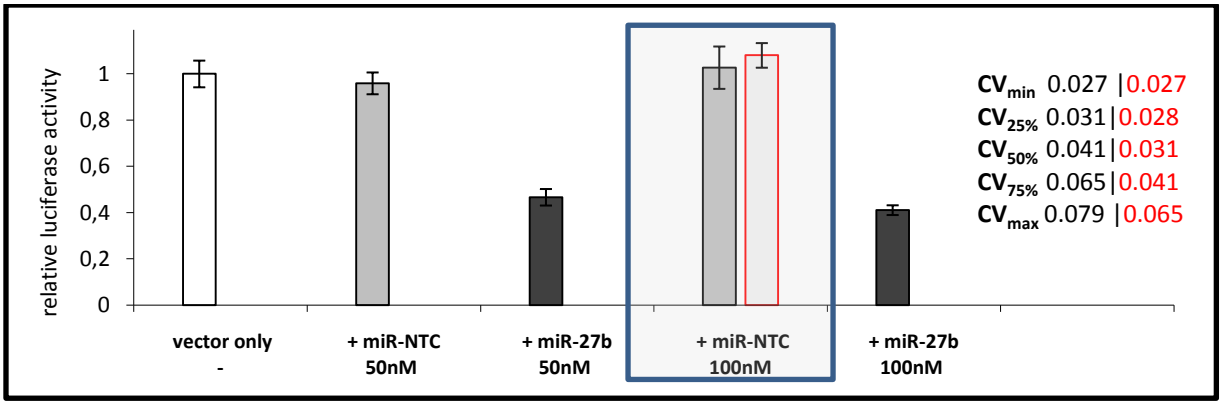


Fig. R39 Influence of PMT-nonlinearity correction on standard deviation and CV values. Shown are the means of the separate technical replicates in **Fig. R38.** In red: mean respectively CV corrected by calculation (verified by repeated measurement of diluted sample).

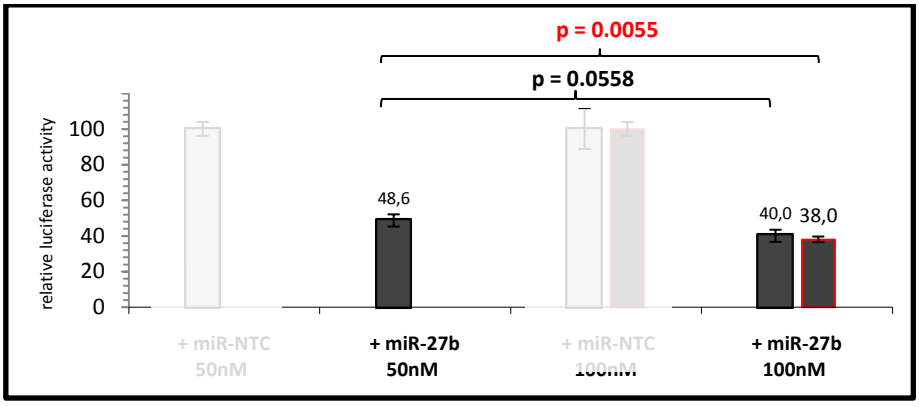


Fig. R40 Influence of PMT-nonlinearity correction on significance tests. In the question whether 100 nM of miRNA mimic would significantly increase the knockdown as compared to 50 nM with respect to the particular non-targeting controls the slightly decreased mean value and the increased variance of the reference value (see mark in **Fig. R39**) caused by PMT-nonlinearity prevents the p-value falling beneath the 5 % alpha level (what should actually happen with respect to the truly existing photons and luminescence reactions, respectively).

The example in **Fig. R40** may make the impression of subtleness, but in cases of rather weak effects (low knockdowns of less than 30 %) and maybe some handling problems of the experimenter causing greater variation that may lead to replicates lying in substantially different raw value ranges, the additional variance caused by this effect may kill the significance of the result. A correction by applying a characteristic curve of the PMT (or rather the luminometer en bloc as also behaviors of some electronic circuits cause photon count drops; personal communication, Berthold) minimizes this influence of the measurement instrument on results – what is *not* done routinely. It is to note that this effect is only to consider when raw values are rather high (above 3 MIO RLUs, see **Fig. R31** and **Fig. R32**) and that it can be avoided or at least minimized by sample dilution, what admittedly may be inconvenient in larger setups. Especially because high value outliers may occur and can naturally only be identified during measurement, thus, repeated measurements are necessary. An alternative would be to apply a characteristic curve of the PMT on the primary data and thereby mathematically correct the values (see **6.1 Outlook**).

4.5 Functionality test – reproducing a published result

Though some results based on different miRNAs and plasmid constructs generated in-house have already been shown above to illustrate the assay behavior, indeed the very first application was the rerun of a published result by *Christoph Gebeshuber* (IMBA, Vienna), who had recently shown that *tristetraprolin* is targeted by miR-29a.

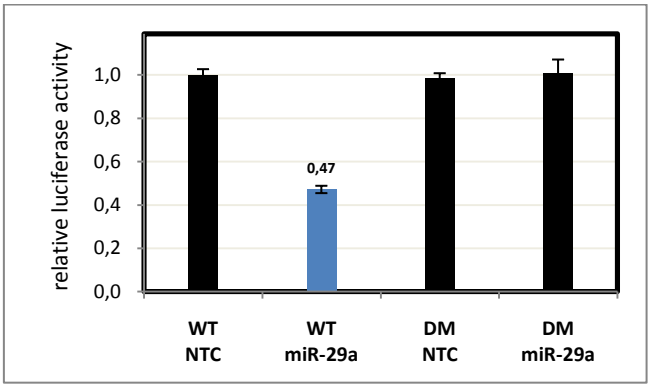


Fig. R41 miR-29a targets TTP (tristetraprolin) – assay result. 18.000 HEK293 cells per well of a 96-well plate were seeded 20 hrs before transfection, 90 ng of FL pGL3 construct (bearing either the two binding sites containing wildtype *WT* or double mutated *DM* form of the TTP 3' UTR) and 10 ng of RL pGL3 reference vector were co-transfected with 100 nM miR-29a or non-targeting control (NTC) miRNA using 0.3 µl DFD. Cells were harvested 48 hrs post transfection. The data shown represent mean±SD from two *biological* (independent) replicates. In a subsequent experiment the knockdown could also be confirmed by CaP transfection (see **Fig. A3**), showing a slightly stronger effect (60 % knockdown) being closer to the published one. TTP-3' UTR constructs were kindly provided by *Christoph Gebeshuber* (IMP, Vienna).

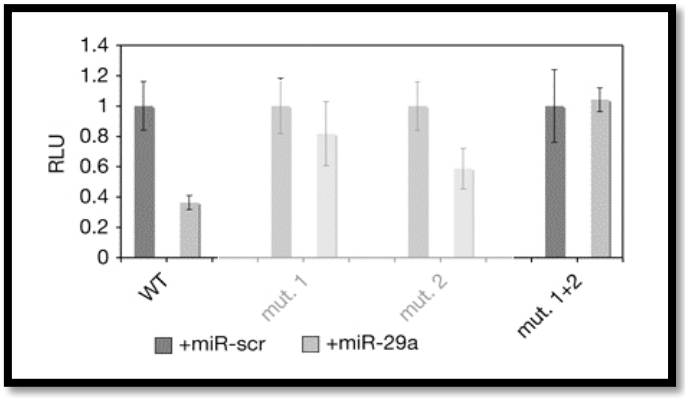


Fig. R42 miR-29a targets TTP (tristetraprolin) – original result. NIH3T3 cells were transfected with miR-scr (scrambled control) or miR-29a. Constructs containing the luciferase ORF coupled to the wildtype (WT) and mutated (mut.) versions of the TTP 3' UTR were used to assess the regulatory impact of miR-29a on TTP. In total, 40.000 NIH3T3 cells were seeded per well of a 24-well plate. 200 ng of pGL3 vector containing the indicated 3' UTR coupled to firefly luciferase plus 100nM miR-29a or scrambled siRNA was transfected per well and 20 ng *Renilla* luciferase was used for normalization. adapted from (Gebeshuber et al. 2009).

4.6 Functionality test – siRNA/miRNA - concentration dependency

RNA transfection reagents such as *DharmaFECT 1-4* (having the reputation of being one of the best on market) typically achieve 90 to nearly 100% siRNA-mediated target knockdown using 100 nM under optimal conditions. Though on first sight the obtained knockdowns may not seem impressing compared to achievable ones with specialized RNA transfection reagents, they indeed are satisfying because the utilized siRNA consisted of a mixture of 4 siRNAs (Dharmacon, *siGENOME* Non-Targeting siRNA Pool #2, Cat#D-001206-14) where only one targets firefly luciferase (siRNA #2, see **Fig. Ax**, Appendix) – noteworthy, with a single mismatch. Moreover, DharmaFECT Duo is formulated to transfect both DNA and RNA at the same time and thus maybe provides non-optimal conditions for RNA transfer. The figure (**Fig. R43**) shows a HEK293 setup, but in HeLa cells the knockdowns were quite similar.

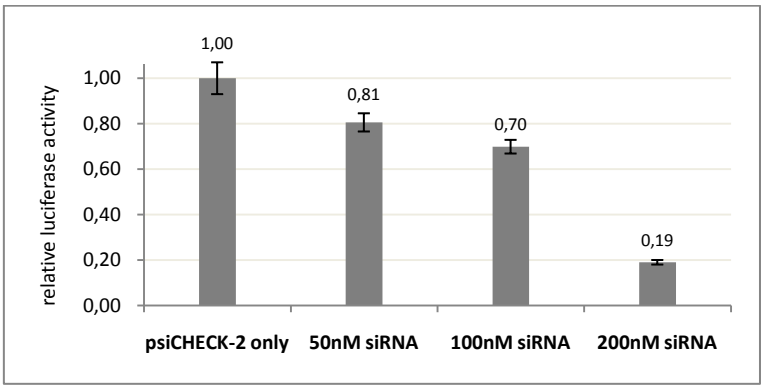


Fig. R43 siRNA dose dependency. *Setup*: 18.000 HEK293 cells seeded 20 hrs prior to transfection, each condition with 100 ng *psiCHECK-2* vector, DFD/NUC ratio ~ 1.5, harvest 48 hrs post transfection.

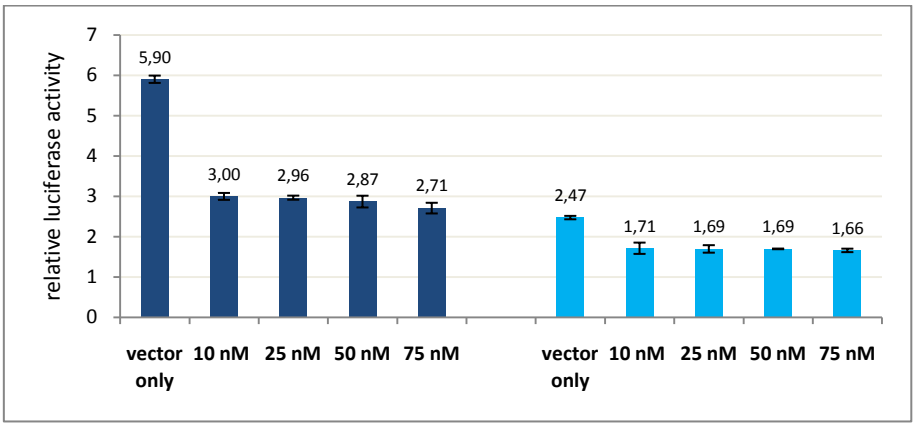


Fig. R44 miRNA mimic dose dependency and effect of endogenous levels. *Setup:* 18.000 cells seeded 20 hrs prior to transfection; 100 ng construct plus miR-27b in indicated amounts; DFD/NUC ratios were kept in a range of 1.6 to 1.8; done without miRNA NTC due to prior results (see **4.7**).

miR-27b was cotransfected with a construct bearing the 3' UTR of PPARγ (see **4.7**) into HEK293 cells (dark blue) or HeLa cells (light blue) (**Fig. R44**). As stated in **4.2** (expression profiling) endogenous miR-27b levels in HeLa cells are about 60% higher as compared with HEK293 cells explaining the generally lower ratios in HeLa cells, and also the less marked knockdown of ~ 30% (in HEK293 ~ 50%).

As in **Fig. R44** (PPARγ and miR-27b), several experiments (conducted by colleagues; e.g. RB1 and miR-26a) indicate that with most miRNA-target construct pairs the concentration dependency in the range below 50 nM is weak or even absent. Also in the examined case of RB1 construct and miR-26a this behavior is evident and only high miRNA concentrations show stronger effects (**Fig. R45** and **Fig. R46**). Though, it is to say that concentrations of 50 nM conform to manufacturer's recommendations and that for assays described in literature typically 50 to 100 nM (or even beyond) of sRNAs were used. In principle, also serial transfections could be conducted. Indeed, small-scale experiments indicate a slight increase in knockdown in treated cases. Serial transfection was not the method of choice mainly due to convenience reasons as even with shortened incubation times of 4 hrs a total time window of 9 hours cannot be undercut. Moreover, the transfection procedure has to be conducted twice, besides loss of time, causing not only more stress on operators but supposably also on cells – where it is to say that effects on viability have not been tested for in these experiments. However, in some important cases which show a rather weak knockdown effect, one may switch to a serial transfection protocol where siRNA/miRNA is transfected first at lower optical confluence and 4 to 6 hrs later the plasmid(s) can follow. Of course, the second transfection can also follow the next day. The established *lipid-nucleic acid* ratios are not necessarily fully optimal for exclusive sRNA transfection but should still yield acceptable results right away.

Although such high sRNA doses are not going to be used in standard assays, the effect of applying 100 nM or 200 nM miRNA were examined for the pair RB1 and miR-26a showing a marked increase of knockdowns. The same transfection media were applied to both cell lines.

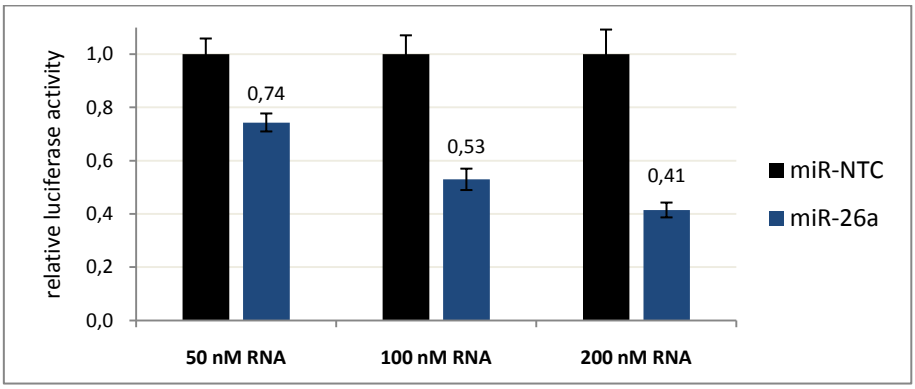


Fig. R45 miRNA mimic dose dependency, miR-26a and RB1 construct – HEK293. Knockdown with respect to appropriate NTC condition increases with miRNA dose. *Setup*: 18.000 cells seeded 20 hrs prior to transfection, time till harvest 24 hrs, 100 ng RB1 construct, either miR-NTC or miR-26a in indicated amounts, DFD/NUC constantly 1.5

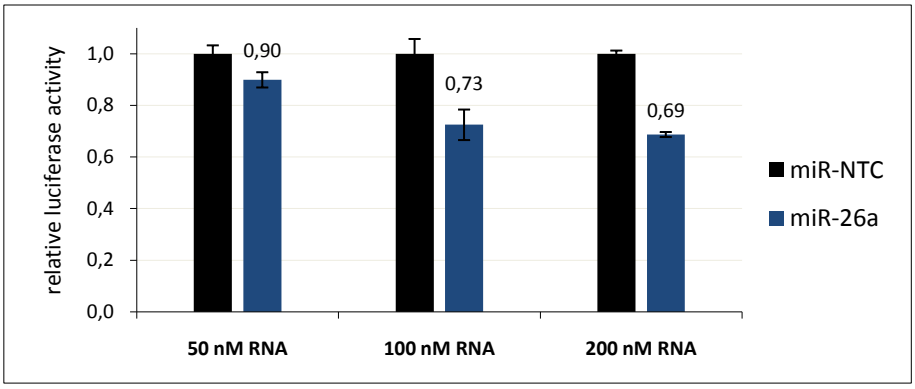


Fig. R46 miRNA mimic dose dependency, miR-26a and RB1 construct – HeLa. Again, knockdown with respect to appropriate NTC condition increases with miRNA dose, but in accordance with other results (**Fig. R25; Fig. R26**) effects are generally lower in HeLa cells (for this pair). *Setup*: 18.000 cells seeded 20 hrs prior to transfection, time till harvest 24 hrs, 100 ng RB1 construct, either miR-NTC or miR-26a in indicated amounts, DFD/NUC constantly 1.5

4.7 Influence of DNA and RNA proportions on results.

Under optimal transfection conditions luminescence raw values (RLU) scale approximately logarithmically with the invested amount of luciferase vector (data not shown). In principle, any amount of 50 to 100 ng is feasible – higher amounts are not necessary (indeed 50 ng already give high values in both cell lines of more than 5 MIO RLU with 5 µl sample measured, depending on the construct) and lower amounts cause a rapid drop in RLUs. Although the pre-tests for FACS analyses (**Fig. R7**) did not indicate significant effects on transfection efficiency when RNA-lacking lipoplexes are used, theoretically different lipoplex-structures may arise when DNA and RNA are present in different proportions, what can have effects on complex moldering and subsequent regulatory functions and transcription, respectively. Thus the following experiments (**Fig. R46** and **Fig. R47**) have been conducted to check for such effects.

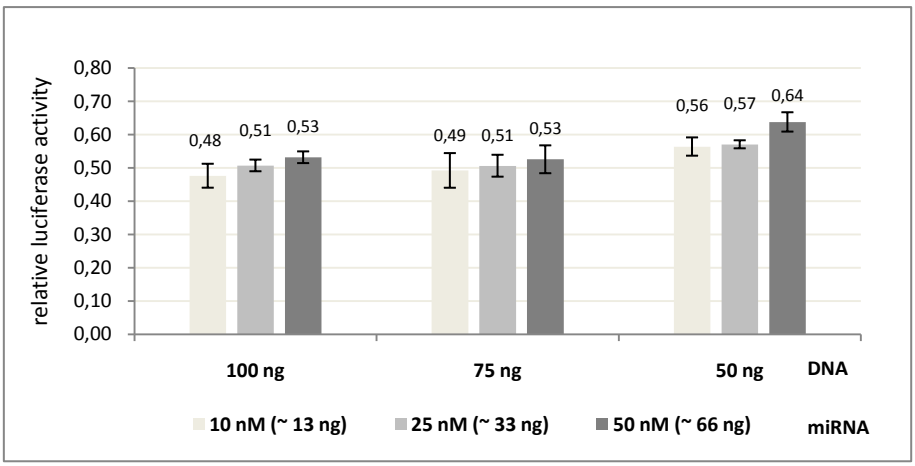


Fig. R47 Effect of lipoplex composition on knockdown – HEK293. *Setup:* RB1 construct transfected with indicated amounts of miR-26a or miR-NTC; DFD/NUC ratios were kept in a range of 1.3 to 1.6; 18.000 HEK293 cells seeded 20 hrs before transfection; 48 hrs incubation until harvest.

For HEK293 cells (**Fig. R47**) there is apparently a trend per block for slightly decreasing knockdowns with **increasing** amounts of miRNA mimic. This is rather contrary to expectation, but has shown up in several experiments in the molarity range below 50 nM (not beyond) what could indicate some transition range of non-optimal DNA/RNA proportions. However, it may depend on the particular plasmid as, for instance, the PPARγ construct (**Fig. R44**) did not exhibit this behavior. Moreover, the differences are not statistically significant. In addition, in the present case knockdowns decrease considerably when low amounts of DNA are used (watch the 50 ng plasmid block, especially the 66 ng miRNA bar). In HeLa cells (**Fig. R48**) there is only an indication of a trend (with 100 ng) for slightly decreasing knockdowns with relatively increasing amounts of miRNA mimic – for this construct.

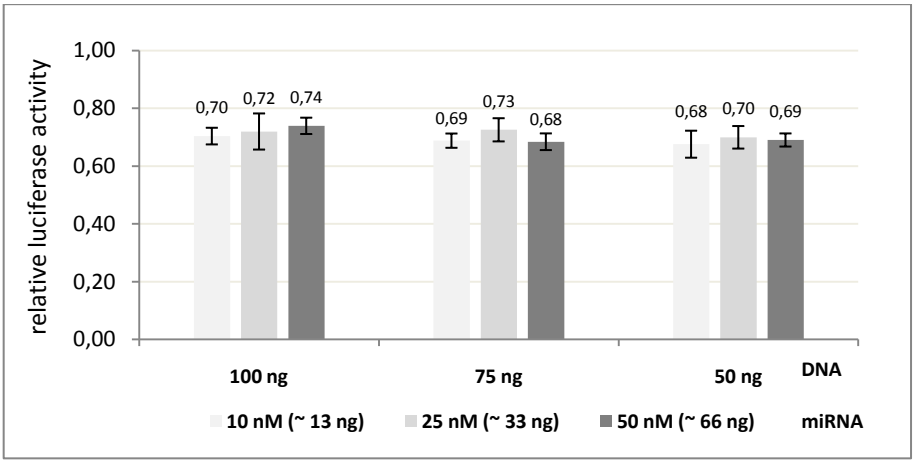


Fig. R48 Effect of lipoplex composition on knockdown – HeLa. *Setup:* RB1 construct transfected with indicated amounts of miR-26a or miR-NTC; DFD/NUC ratios were kept in a range of 1.3 to 1.6; 18.000 HeLa cells seeded 20 hrs before transfection; 48 hrs incubation until harvest

Both cell lines have been transfected with the same transfection media and measured at same day with equal sample preparation and the same assay reagents, thus the differences in knockdowns and trends are clearly cell line specific. As all differences are non-significant it is no major point to think of in planning an assay experiment, but at least with HEK293 cells it may be better not to scale down DNA amount below 75 ng and to make a small pre-test with 10, 25 and 50 nM for a particular construct if maximal knockdowns are required.

4.8 First application – PPARγ and miR-27b

The first scientifically relevant application conducted in-house was the validation of the predicted but not experimentally validated interaction pair PPARγ and miR-27b. Experiments in human preadipocytes demonstrated that miRNAs are involved in adipogenesis by impairment of adipocyte differentiation upon miR-143 inhibition (Esau et al. 2004). Furthermore, Klöting et al. identified differentially expressed miRNAs between omental and subcutaneous fat (Klöting et al. 2009), indicating that further miRNAs might regulate adipogenesis in human. In-house conducted *qPCR* and *microarray* analyses showed decreased expression of miR-27b during adipogenesis of human adipose-derived stem (hMADS) cells (40% as compared with proliferating cells), and overexpression of miR-27b blunted early induction of PPARγ and C/EBPα and repressed adipogenic marker gene expression and triglyceride accumulation in late stages. Based on the previously in-house constructed prediction data set (**1.6**) PPARγ was revealed as putative miR-27b target with a highly conserved binding-site in its 3' UTR, consistently *qPCR* results exhibit anticorrelated trends of expression levels during differentiation (**Fig. R49**) (Karbiener et al. 2009).

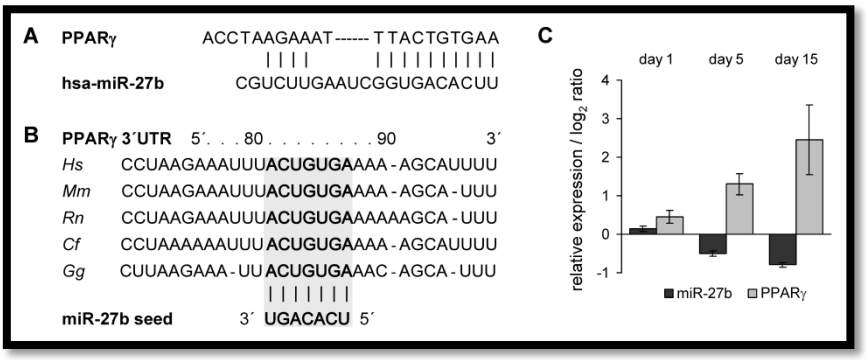


Fig. R49 Bioinformatic results and supporting experimental evidence for PPARγ 3' UTR miR-27b interaction. Indeed, the binding site prediction (A) combined with the strong conservation (B) indicates a reliable prediction. However, startling many high-confidence predictions fail the experimental validation in the end. Reciprocal expression levels (relative to day 0 of differentiation) determined by *qPCR* experiments (C) also indicate a relation but not necessarily a direct interaction. adapted from (Karbiener et al. 2009)

The luciferase assay result (**Fig. R50**) demonstrates the functionality of the predicted miR-27b response element and together with expression analysis results strongly indicates that miR-27b functions as repressor of human adipogenesis by directly targeting PPARγ. For this assay the 3' UTR of PPARγ2 (NM_0156989; **Tab. A3**, Appendix) was cloned into the *psiCHECK-2* vector downstream of the *Renilla* luciferase coding sequence using the *Not I* and *Xho I* restriction sites in the MCS (see Methods). To monitor specificity a second plasmid with 4 mutated bases in the seed sequence was constructed (for details on sequences see **Tab. A3**, Appendix). Assay data clearly demonstrate that miR-27b specifically interacts with the predicted binding-site in PPARγ's 3' UTR.

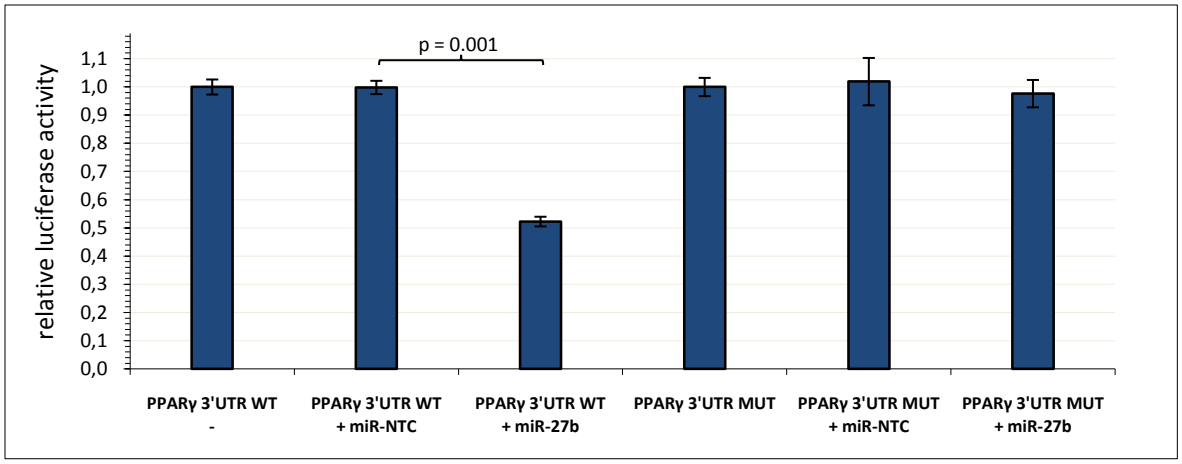


Fig. R50 miR-27b targets PPARγ. *Renilla* luciferase activity (RLU) was normalized to firefly luciferase activity first and all conditions were normalized to the respective *vector only* transfections. The wildtype (WT) construct in interaction with miR-27b clearly shows suppression of *Renilla* activity (by ~ 50 %) as compared with vector only and non-targeting controls (miR-NTC). Moreover, introduction of a 4 base mutation in the seed sequence (MUT construct) completely rescues *Renilla* expression. For details on sequences see **Tab. A3** (Appendix). *Setup:* 20.000 HEK293 cells were seeded 18 hrs prior to transfection of 100 ng construct bearing either the wildtype (WT) or mutated (MUT) version of the 3' UTR of PPARγ and either 50 nM of miR-27b or miR-NTC, or neither; DFD/NUC ratio ~1.2

As additional evidence and as test for inhibitor transfection and action, in a subsequent experiment an antisense oligonucleotide (*ASO*) targeting miR-27b (by hybridization) has been transfected in combination with miR-27b (**Fig. R51**). The fact that the level of the *ASO* condition reaches far beyond the vector only value may be assigned to the existence of endogenous miR-27b (proven by miRNA profiling) also being inhibited – as outlined in previous results endogenous levels decrease basal *Renilla* levels.

.

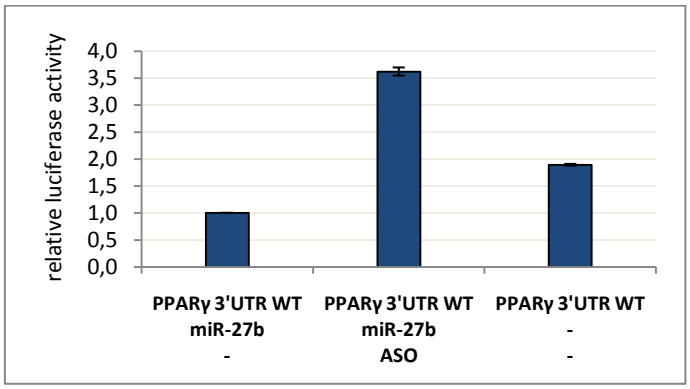


Fig. R51 Effect of antisense oligonucleotides (*microRNA Inhibitors*). *Setup:* 20.000 HEK293 cells were seeded 18 hrs prior to transfection of 75 ng construct bearing the wildtype 3' UTR of PPARγ and either 50 nM of miR-27b alone, 50 nM miR-27b in combination with 50 nM *ASO*, or neither; DFD/NUC ratio ~1.4

5. Discussion

5.1 Discussion of material & methods

5.1.1 Cell culture and preparation for transfections

As cells provide the chemical environment and especially the protein machines that actually perform the core work of the assay, they must be treated tenderly. Data indicate that despite the known reactivity of cellular processes on environmental stimuli the variance introduced thereby can be held minimal when treatment is strictly kept constant – this especially concerns temperatures, trypsinization procedures and seed counts/incubation time. Temperature and above all timings are also essential for achieving constant transfection results – reagents must be at room temperature before usage and incubation times according to the protocol must be adhered to or at least be kept constant between repeated experiments.

5.1.2 FACS analysis

5.1.2.1 Determination of transfection efficiency

GFP needs rather strong promoters to drive sufficient expression for detection, especially in mammalian cells. Most published examples (also those using brightened GFPs with mutations to promote folding at 37°C) have used constitutive promoters from viruses such as cytomegalovirus (CMV), simian virus 40 (SV40). Indeed, it has been estimated that 1 µM well-folded wild-type GFP molecules are required to equal the endogenous autofluorescence of a typical mammalian cell. Enhanced GFP variants (as used in the present study) with improved extinction coefficients might improve the detection limit up to tenfold, but 0.1 µM GFP is still approximately 10^5 copies per typical cell of 1 to 2 pl volume. This estimate already assumes perfect GFP maturation. Thus sensitivity limit is not set by instrumentation but by cellular autofluorescence (Tsien 1998). Moreover, studies indicate that GFP expression not necessarily reflects all successfully transfected cells. Tseng et al. have shown that although 95% of treated cells contained the plasmid after 3 hrs of exposure to a liposomal vector, transgene expressing cell proportion achieved its maximum of only 30% (!) after 48 hrs. These results were consistent with previous evidence indicating that the transmembrane transport of cationic liposomal vectors is *not* the limiting step in transgene expression. Apparently plasmid copy number per transfected cell is very important, where conditions which permit the plasmid entry flux to exceed the intracellular DNA degradation rate presumable enhance transgene expression (Tseng et al. 1997).

Thus, true GFP^+ rates and especially proportions of in fact transfected cells, respectively, are almost certainly higher than stated, and may also explain some inconsistency in FACS and luminometer data. However, the true viability of cells and the relative best transfection condition with respect to nucleic acid uptake could be determined with this method.

5.1.2.2 Determination of viability

Assessing viability of adherent cells using FACS and dyes only capable to penetrate membranes with in a way disturbed integrity is not an easy task, as detaching of cells by trypsinization and subsequent separation of cell clumps by pipetting (what is absolutely necessary in FACS analysis even though duplets may be considered in data analysis) introduce membrane damage in any case. However, untreated cells (being used for viability normalization) showed 7-AAD positive rates of maximally 10% (prolonged time till measurement; 2.5 hrs), and in most cases below 5%, thus the established sample preparation protocol is acceptable.

5.1.2 Transfection

5.1.2.1 Need for optimizations of cell culture and transfection conditions

Endocytosis and proliferation properties of cells play a major role in lipid-based (generally synthetic carrier based) transfection methods as lipoplexes pass the membrane by endocytosis, and decomplexed DNA can (in most cases) only reach the transcription machinery in the nucleus during cell division, when the nuclear membrane is decomposed (Kreiss et al. 1999). Thus, it is necessary to maintain optimal culture conditions and seed cells at a density that ensure high proliferation rates and normal membrane processes at the time of transfection. Different membrane properties (especially with respect to charge density) and also specific intracellular conditions (especially nuclease activity) are two major factors because of which transfection optimization has to be performed for each cell line – even when the same transfection reagent is used. Complicating, the stated influencing variables also differ between different genotypes of the same cell line (e.g. HeLa populations from different labs) and may also be observable after long-term passaging of a particular cell line (that is why I would recommend to use only the first 5 passages for assays as there are plenty of cryo stocks available). Clearly, an established protocol for a specific transfection reagent (e.g. DFD) cannot simply be used for another one (e.g. MP), as each reagent has its characteristic composition (e.g. amounts of cationic lipids and colipids), lipid structure (e.g. liposome building or not) and physicochemical properties determining lipoplex structure and charge (Promega, Biontex).

5.1.2.2 Critics

Though, lipid-mediated transfection is widely used for siRNAs and miRNAs, Barreau et al. have pointed out that liposome-based RNA transfection may face the problem of protecting compartment formation (RNA-associated fluorescence was concentrated in punctuate structures) what led to abnormal stability of mRNAs in the cited study, likely because molecules were sequestered away from the cytoplasmic machineries responsible for mRNA degradation (and probably also translation). However, some transfected RNAs must have been released into the cytoplasm as the production of a fluorescent reporter they were coding for could be observed (Barreau et al. 2006). These results may explain why rather high quantities of sRNAs (> 50 nM) were necessary to markedly improve knockdowns in the present work, where the problem arises that high amounts of unused transfected sRNAs may interfere with observed effects or become toxic to cells and trigger a response such as stress causing a change in cell metabolism – though, there are recent developments showing

improvement (Kim et al. 2010, Damen et al. 2010, Wang et al. 2010). As to that, the use of plasmids encoding *shRNAs* (small hairpin RNAs) expressing siRNA or miRNA precursors may be preferable to transfection of the chemically synthesized mature versions utilized. Thereby, knockdowns might be enhanced substantially, but the price is the need for a cloning procedure for each sRNA that shall be examined in its effect.

Also lipid-based DNA transfection faces inefficiency where (as with RNA) it is typically *not* the inability to enter the cells, as several studies have shown that most cells in a transfection take up the DNA (Tseng et al. 1997, Zabner et al. 1995). Part of inefficiency may be caused by the fact that the population of lipid-DNA complexes is very heterogeneous, even under conditions optimized to produce best transfection results. After endocytosis, lipid-DNA may aggregate into large perinuclear complexes often showing a highly ordered tubular structure, releasing only a smaller amount of free DNA into the cytoplasm. Furthermore, for transcription DNA not only has to get away from lipids, it also has to pass the nuclear membrane (Zabner et al. 1995).

Co-transfection approaches are based on an all-or-none assumption – being rarely verified experimentally – that any cell takes up all sorts of transfected molecules (e.g. two different plasmids, or sRNA and DNA) proportionately to their concentrations present during transfection. This assumption is also made in the typical setup of co-transecting a plasmid coding for a fluorescent marker protein (e.g. GFP) in order to distinguish transfected from untransfected cells. Indeed, co-transfections may not work in that simple way. Experiments conducted by Ma et al. have shown a maximal co-transfection efficiency (expression of both eGFP and a second protein per cell) of ~ 50% (Ma et al. 2007). Thus, in the present work an additional confocal z-section scanning of cells transfected with a fluorescent reporter protein and a fluorescence-labeled siRNA would have been interesting.

Even if all plasmids in DNA co-transfections enter a cell there is the additional problem of *trans* effects between promoters that can potentially affect reporter gene expression and thereby compromise results. This point is irrelevant using the *psiCHECK-2* vector but may become important when *pGL4* vectors are used for promoter studies.

5.1.2.3 Decision

There are plenty of transfection reagents on the market, and it is definitely possible that another one would perform better – in the sense of stronger miRNA-related effect due to more efficient RNA transfection, for instance. But, DharmaFECT Duo yields good effects (50% knockdown with a single 3' UTR binding site and 50 nM RNA; what is in the typical maximum range of strength of effect) and above all very stable and reliable results. These facts combined with its easy handling and a convenient protocol (as there is no need to remove transfection complexes within 72 hrs of incubation) makes it a good choice. METAFECTENE Pro also yields good results (indeed, slightly stronger effects with siRNA and maybe miRNA), but with much higher variation with corresponding negative effects on statistical tests and on top of that the need for complex removal within 6 hrs introducing an extra protocol step. Variability of results with respect to biological replicates has also been the main point against the *calcium phosphate* method. In principle, transfections could also be performed in-house per electroporation with a *Neon* Transfection System (*Invitrogen*), what might yield the best results (also with respect to viability, typically the weak point of electroporation but *Invitrogen* seems to have improved the technology substantially). But, as the per-sample-price

(assuming triplicates) is approximately 20 times higher as compared with lipid-based methods this would be economic nonsense.

5.1.3 Luciferase vector

The *psiCHECK-2* vector performs very well. Meanwhile a special *pmirGLO* (*Promega,* Cat#E1330) is available bearing also *Renilla* and firefly luciferases but under the control of a weaker human phosphoglycerate kinase (PGK) promoter, what is likely to make the assay more sensitive and possibly sample dilution unnecessary – but, this vector change may call for some minor transfection adjustments. A more interesting substitution concerns *pGL4* vectors (*Promega*). These luciferase vectors are suited for promoter studies, extending the usability of the established assay (*pGL4* vectors are in-house and work well under both DharmaFECT Duo and METAFECTENE Pro transfections based on conditions stated above; data not shown).

5.2 Discussion of results

5.2.1 Technical issues

5.2.1.1 Timing and RNA degradation

In general, RNAs have inherent half-lifes determining their level of accumulation, where the stability correlates strongest with structural features, but can be altered in response to external stimuli such as hormones or various types of stress. In the case of mRNAs life time is typically in the order of several hours (median estimated half life of 7.1 h based on nearly 20.000 genes in mouse ES cells; less than 100 genes showed a mRNA half life of less than 1 h), where short-lived are enriched among genes with regulatory functions (such as transcription factors), whereas mRNAs with long half-life are enriched among genes related to metabolism and structure (cytoskeleton, extracellular matrix). mRNA stability correlates most significantly positive with the number of exon junctions per open reading frame length, and negative with the presence of PUF-binding motifs, AU-rich elements in 3' UTRs and CpG dinucleotides in 5' UTRs (Sharova et al. 2009). In a time course study using HEK293 cells, conducted by Bail et al., *actinomycin D*-directed transcriptional shutoff revealed that most of the traced endogenous miRNAs (95% in a microarray experiment, but subsequent qPCR of selected miRNAs indicate a yet higher proportion) remained stable and did not significantly change throughout 8 hrs, but in a few cases shorter half lifes could be observed. For instance, miR-382 levels were reduced more than 50% within this time window (Bail et al. 2010).

With regard to *luc*-fusion mRNA life-time these results may be irrelevant, as time course experiments (see **Fig. R27** and **Fig. R28**) indicate that protein amount increases in both HEK293 and HeLa cells throughout 48 hrs of incubation (the typical duration) – though, stress-induced effects on *luc* mRNA life-time cannot be ruled out and may explain some variation in biological replicates, furthermore the conducted experiment is not universally valid and is expected to draw a different picture with a different *luc*-fusion construct used (indeed, in the utilized RB1 construct the *luc* mRNA is generously

extended by the rather long RB1 3' UTR, what might prolong life-time of the *luc*). The data on miRNA half-lifes on the other hand may be of importance in rare cases, as a level-drop of, for instance, 50% within the first 8 hrs would certainly impair the measured knockdown, especially in the typical setup with a 48 hrs incubation time. But, one has to be aware of the fact that transcriptional shutoff is used to set the zero point and trigger dynamics in half-life studies, what in the case of miRNAs means the observation of the whole biogenesis pathway with according uncertainty at which point their lifes mainly end – as mentioned above (**1.4.2 microRNA Biogenesis**) mature miRNA duplexes seem to be very stable, and that is the structure transfected *miRNA mimics* have. Moreover, time course experiments (see **Fig. R25** and **Fig. R26**) have shown that knockdown may even increase during 72 hrs of post transfection incubation – this is not necessarily true for all *luc* construct-miRNA pairs and should be determined in a pre-test in each case, if maximal knockdown is strongly favored.

5.2.1.2 Target design

As mentioned in **1.4.3 microRNA Function**, some miRNAs have binding sites in both UTRs of their target mRNA. Therefore, in order to fully understand miRNA function and reveal the true strength of effects, it would be advisable to insert both 5' and 3' UTR sequences in miRNA functional experiments, what has rarely been done so far. Experiments using 3' UTRs alone usually achieve ~ 40% – 60% (max) protein reduction. According to results from Lee et al. more protein reduction may be seen with 5'-UTR inclusion – where interaction sites exist (Lee et al. 2009). Though, for convenience this would need a special cloning vector with an additional 5' MCS.

Furthermore, for improved results alternative 3' UTRs have to be considered as far as possible based on the present annotations.

Although it is not uncommon to insert several copies of a target site to enhance efficacy and boost knockdown, this approach does certainly not simulate *in vivo* conditions and results may be questionable, especially with respect to seed context and target-site accessibility (both important as mentioned above). Thus, all in-house created target constructs bear the whole unmodified 3' UTR of the predicted target mRNA (only truncated when cloning procedure requires it) in order to provide as natural conditions as possible (but of course the action takes place in a model cell system).

5.2.1.3 Influence of temperature in luminescence measurements

In chemical reactions temperature is a major pace maker. As the measured light intensity in luciferase assays is a measure of the rate of catalyses by luciferases it also depends on temperature, where the optimum for luciferase activity is approximately room temperature (20 to 25 °C). Thus it is important that the assay reagents are fully equilibrated to ambient temperature before beginning measurements – leave them in a water bath maintained at ambient temperature and equilibrate for 30 min (do *not* go beyond 25 °C). Although the luciferase activity is stable for several hours at room temperature in 1x PLB, samples may be left on ice for up to 12 hrs. Assaying cold samples results in a decrease of enzyme activity thus also the samples should be at ambient temperature when measurement begins.

Assay reagents may be stored in aliquots at -20°C for a month or at -70°C for up to 1 year after reconstitution or initial use. After preparation or thaw, reagents should be mixed well before use. Generally, substrates and buffers should be stored in the dark and also be protected from prolonged exposure to light ahead and during measurements (Promega Technical Bulletin #281).

5.2.1.4 General recommendations

Keep the handling procedures as constant as possible from the very beginning until the end of the assay. Always mix all reagents gently and let them fully equilibrate to room temperature before use. Do a pre-test in both cell lines (or based on expression profiling on just one) for all constructs subsequently assayed, determining the optimal incubation time until harvest (with transfections of the respective miR-candidate and miR-NTC to determine the point in time of maximal knockdown). For the optimal incubation time do a second pre-test in the chosen cell line with different amounts of miRNA (e.g. 10 nM, 50 nM and 100 nM) to optimize knockdown. With respect to assay stability the use of technical replicates is not necessary in pre-tests, thus material can be saved. What was said also holds true for inhibitor-experiments.

5.2.2 PPARγ and miR-27b

PPARγ is a factor of wider importance. For instance, research in the last decade has uncovered the presence of all PPARs in most brain cell types, and has shown that their activation, particularly that of PPARγ, is implicated in normal brain and cerebrovascular physiology, and confers protection under pathological conditions. Synthetic ligands for PPARγ (*thiazolidinediones*, TZDs; also known as *glitazones*) are currently only used as medication in the treatment of type 2 diabetes, but accumulating evidence is highlighting the therapeutic potential of PPARγ ligands in the treatment of brain disorders such as *Alzheimer's* disease, where clinical trials with *pioglitazone* and *rosiglitazone* seem promising (Nicolakakis & Hamel 2010). Moreover, also miR-27b is expressed in neurons (Sempere et al. 2004) – probably a thing worth checking into.

PPARγ is also involved in *atherosclerosis* and inflammation, additionally it is expressed in blood cells and induced during macrophage differentiation, for instance (Lehrke & Lazar 2005).

You are what you eat. This saying seems immediately obvious in terms of matter, but it also refers to the functional level. However, things are not that easy down there – as usual in biological systems. Food products may contain rather beneficial or rather adverse molecules, and rather often certainly both, and sometimes the mixture of molecules within the particular product or the meal as a whole decides whether it is a good or a bad one. As an example for the typically good ones, PPARγ may be activated by a number of phytochemicals. Khateeb et al. have shown that polyphenols present in *pomegranate* activate PPARγ via a cAMP-PKA-PPARγ signaling cascade (cAMP, *cyclic adenosine monophosphate*; PKA, *protein kinase A*), which in turn activates serum *paraoxonase 1* (PON1) gene expression. PON1 is synthesized and secreted by the liver and has the capability to protect LDL and HDL from oxidation, to decrease macrophage oxidative status, to decrease oxidative status in atherosclerotic lesions, and to attenuate atherosclerosis development. Interestingly, it has also been

shown that inflammatory factors such as *interleukin 1* and *6*, oxidized phospholipids, and *tumor necrosis factor-α* (look back at what stressed adipocytes send off) decrease PON1 mRNA levels (Khateeb et al. 2010, Feingold et al. 1998). Relationships of that sort might be the reasons why an apple a day keeps the doctor away.

In line with expectations regarding microRNAs also the examined miR-27b is involved in several regulatory pathways and diseases. Besides its role in adipose tissue presented here, and its potential role in neurons or *Alzheimer's* disease as speculated above, it has been shown to play a role in osteoblasts differentiation (reporting miR-27) (Wang & Xu 2010). Furthermore, its levels are significantly increased in sclerotic tissue samples and serum of patients suffering from *arteriosclerosis obliterans* (ASO), a kind of peripheral arterial disease (Li et al. 2010). miR-27b has also been shown (by *luc* assays) to directly target the *matrix metalloproteinase 13* (MMP-13) in human chondrocytes (Akhtar et al. 2010) and it seems to be involved in the post-transcriptional regulation of *adenosine 2B receptor* (Kolachala et al. 2010) – also *luc* validated. Finally, the result of the present work is supported by the finding of Jennewein et al. who also found (in the context of inflammation) that miR-27b targets PPARγ's 3' UTR (Jennewein et al. 2010).

From the technical perspective the achieved knockdown of 50 % in relative luciferase activity is satisfying as according to Baek et. al 'even an overexpressed miRNA typically downregulates most of its endogenous targets by less than 50%' (Baek et al. 2008).

5.3 Recent developments in target identification

Until recently miRNA target identification has relied mainly on computational approaches since genome-wide experimental strategies have been limited by technological development. However, the progresses in high-throughput sequencing did not only have dramatical influence on miRNA gene finding, but also on miRNA target prediction – and certainly likewise on studying gene expression, DNA copy number variations, polymorphisms and posttranscriptional modifications in future.

HITS-CLIP (high-throughput sequencing of RNAs isolated by crosslinking immunoprecipitation) (Licatalosi et al. 2008) is a new method based on purifying RNA binding proteins after UV-crosslinking with their binding partner – enabling large scale experimental data collecting relevant to miRNA target binding in an appropriate setup. Chi et al. used HITS-CLIP to covalently crosslink Argonaute protein-RNA complexes. High-throughput RNA sequencing following purification and partial RNA digestion generated two data sets – Ago-miRNA and Ago-mRNA binding sites – that were combined with bioinformatic analysis to identify miRNA-target mRNA interaction sites (Chi et al. 2009). *Ago HITS-CLIP* seems to outperform bioinformatic predictions alone as the estimated specificity (~ 93 %), false positive (~13-27 %) and false negative (~15-25 %) rates show a major improvement compared with previous computational results (~up to 66 % false positive and ~50-70% false negative rates (Baek et al. 2008, Selbach et al. 2008, Easow et al. 2007)).

However, experimental validation of targets is still indispensable and requires some kind of reporter assay.

5.4 Luciferase assays in general – criticism and why things are not that bad

Although experimental validation of miRNA targets is currently commonly done through *in vitro luciferase assays* which certainly provide some measure as to whether a miRNA binds to an expected target site, one concern with this type of assay is that a miRNA-target pair validated *in vitro* is not necessarily relevant *in vivo*. In addition, typical assay setups do not take into account the possibility of a combinatorial nature of miRNA regulation and rather concentrate on a single miRNA-target pair – thus, an observed missing or non-significant down-regulation of the target transcript might cause the erroneously rejection of the pair investigated. Moreover, *luciferase assays* are relatively expensive and time-consuming to conduct – another reason for the limited amounts of validated targets, as high-throughput is not feasible.

Nevertheless, there is no alternative available or even within sight so far. But, indeed, placing the *luciferase assay* in the right context of accompanying methods such as mRNA and miRNA *cDNA microarray* expression profiling, and *qPCR* analysis – appropriately used enhancing *in vivo* relevance, and the use of natural full 3' UTRs together with co-transfections of several miRNAs (of course, meaningfully using only high confidence predictions) may lower false rejections. The economic issues can be counteracted somewhat by right choices of materials and optimization of their usage. For instance, *DharmaFECT Duo* in combination with the microplate luminometer *ORION II* allow for a total assay time of 3 to 4 days (total work time in dependence of samples 3 to 6 hrs) in only 3 sessions (cell seed, transfection, harvest and measurement). Furthermore, optimizations on several points in the assay enable amount reductions and hence hold down costs.

5.5 In a nutshell – why research on microRNAs makes sense and how to do it currently

Differential miRNA and mRNA expression in cellular processes such as differentiation, proliferation or apoptosis is an important fact that may allow for determination of disease associated genes that are specifically regulated by miRNAs. However, in most cases miRNA target genes are unknown, and determination of changes in miRNA expression can only serve as starting point in the characterization of physiological and pathological processes. These results need to be complemented with bioinformatic predictions or experimental methods such as HITS-CLIP to identify putative miRNA target genes followed by the functional validation where the results of the present work come into play. Thereby relevant insight into cellular regulatory networks can be gained and novel clinically utilizable targets and (sRNA-)drugs may be identified and developed. Very interesting in that context is the finding that microvesicles containing RNA (~ 7000 mRNAs and 140 microRNAs in the cited study) are released from adipocytes, known as adipocyte-derived microvesicles (ADMs), mediating RNA transport to distant cells such as macrophages (Ogawa et al. 2010).

6. Outlook

6.1 Non-linearity of the PMT

As mentioned above, non-linearity of the PMT/luminometer has adverse effects on results, if measurements take place in a too high raw value range. Although the distribution of measurement points in **Fig. D1** is not ideal, it can be seen as first hint of how dramatic the effect is – indeed, the fit was nearly confirmed by a second experiment spanning the range up to 14 MIO RLU ($y = -1E\text{-}08x^2 + 0.91x$; $R^2 = 0.9991$). It is recommendable to measure at least 10 linearity curves (using at least serial dilutions of samples with high luciferase concentration, but better a defined standard) with a closer spacing of measurement points in the higher raw value range above 5 MIO RLU. Combining these results to a calibration curve, this data may be used to automatically correct RLU raw values for deviations purely introduced by the measurement device. Indeed, a short *R* or *MATLAB* script being able to load the *Excel* file exported by *Simplicity*, to extract the raw values, to apply the transformation and to write back the corrected values would suffice.

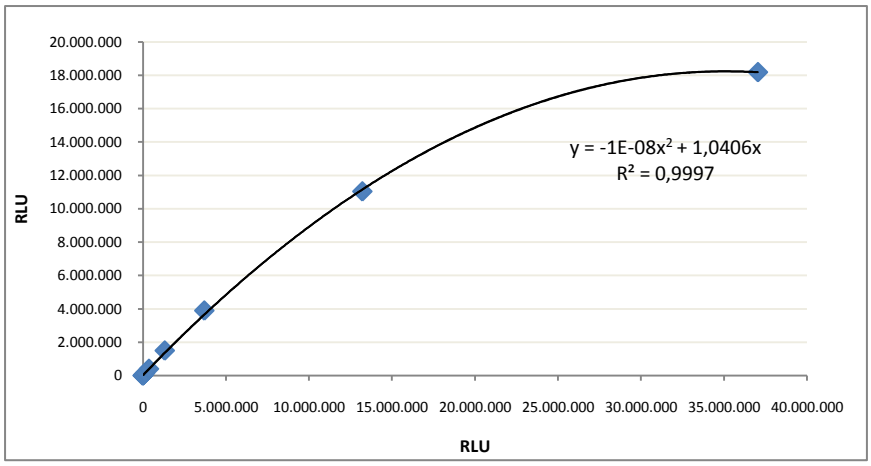


Fig. D1 Non-linearity of the PMT may influence result interpretation. On the x-axis there are the 'true' RLU values inferred from low signal values (lying in the linear range) and dilution factors, on the y-axis there are the in fact measured values.

7. References

ENCODE 2007. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature*, 447(7146), pp.799-816.

Aalberts, D.P. & Nandagopal, N., 2010. A two-length-scale polymer theory for RNA loop free energies and helix stacking. *RNA*. Available at: http://rnajournal.cshlp.org/content/early/2010/05/24/rna.1831710.abstract.

Addo-Quaye, C., Miller, W. & Axtell, M.J., 2009. CleaveLand: a pipeline for using degradome data to find cleaved small RNA targets. *Bioinformatics*, 25(1), pp.130-131.

Ahmad, A. et al., 2005. New multivalent cationic lipids reveal bell curve for transfection efficiency versus membrane charge density: lipid-DNA complexes for gene delivery. *The Journal of Gene Medicine*, 7(6), pp.739-748.

Akhtar, N. et al., 2010. MicroRNA-27b regulates the expression of matrix metalloproteinase 13 in human osteoarthritis chondrocytes. *Arthritis and Rheumatism*, 62(5), pp.1361-1371.

Altuvia, Y. et al., 2005. Clustering and conservation patterns of human microRNAs. *Nucleic Acids Research*, 33(8), pp.2697-2706.

Antuna-Puente, B. et al., 2008. Adipokines: the missing link between insulin resistance and obesity. *Diabetes & Metabolism*, 34(1), pp.2-11.

Arnosti, D.N. & Kulkarni, M.M., 2005. Transcriptional enhancers: Intelligent enhanceosomes or flexible billboards? *Journal of Cellular Biochemistry*, 94(5), pp.890-898.

Azuma-Mukai, A. et al., 2008. Characterization of endogenous human Argonautes and their miRNA partners in RNA silencing. *Proceedings of the National Academy of Sciences of the United States of America*, 105(23), pp.7964-7969.

Baek, D. et al., 2008. The impact of microRNAs on protein output. *Nature*, 455(7209), pp.64-71.

Bail, S. et al., 2010. Differential regulation of microRNA stability. *RNA (New York, N.Y.)*. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20348442 [Accessed April 1, 2010].

Bailey, A.S. et al., 2006. Myeloid lineage progenitors give rise to vascular endothelium. *Proceedings of the National Academy of Sciences of the United States of America*, 103(35), pp.13156-13161.

Bamshad, M., Song, C.K. & Bartness, T.J., 1999. CNS origins of the sympathetic nervous system outflow to brown adipose tissue. *Am J Physiol Regul Integr Comp Physiol*, 276(6), pp.R1569-1578.

Barbatelli, G. et al., 2010. The emergence of cold-induced brown adipocytes in mouse white fat depots is determined predominantly by white to brown adipocyte transdifferentiation. *American Journal of Physiology. Endocrinology and Metabolism*, 298(6), pp.E1244-1253.

Barreau, C. et al., 2006. Liposome-mediated RNA transfection should be used with caution. *RNA*, 12(10), pp.1790-1793.

Bartel, D., 2004. MicroRNAsGenomics, Biogenesis, Mechanism, and Function. *Cell*, 116(2), pp.281-297.

Bartel, D.P., 2009. MicroRNAs: Target Recognition and Regulatory Functions. , 136(2), pp.215-233.

Baskerville, S. & Bartel, D.P., 2005a. Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. *RNA (New York, N.Y.)*, 11(3), pp.241-247.

Baskerville, S. & Bartel, D.P., 2005b. Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. *RNA (New York, N.Y.)*, 11(3), pp.241-247.

Behm-Ansmant, I. et al., 2006. mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. *Genes & Development*, 20(14), pp.1885-1898.

Benes, V. & Castoldi, M., 2010. Expression profiling of microRNA using real-time quantitative PCR, how to use it and what is available. *Methods*, 50(4), pp.244-249.

Berezikov, E. et al., 2007. Mammalian mirtron genes. *Molecular Cell*, 28(2), pp.328-336.

Berezikov, E. et al., 2005. Phylogenetic Shadowing and Computational Identification of Human microRNA Genes. , 120(1), pp.21-24.

Berezikov, E. et al., 2006. Diversity of microRNAs in human and chimpanzee brain. *Nature Genetics*, 38(12), pp.1375-1377.

van den Berg, A., Mols, J. & Han, J., 2008. RISC-target interaction: cleavage and translational suppression. *Biochimica Et Biophysica Acta*, 1779(11), pp.668-677.

Bernstein, E. et al., 2001. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature*, 409(6818), pp.363-366.

Bertram, J., 2006. MATra - Magnet Assisted Transfection: combining nanotechnology and magnetic forces to improve intracellular delivery

of nucleic acids. *Current Pharmaceutical Biotechnology*, 7(4), pp.277-285.

Bhattacharyya, S.N. et al., 2006. Stress-induced reversal of microRNA repression and mRNA P-body localization in human cells. *Cold Spring Harbor Symposia on Quantitative Biology*, 71, pp.513-521.

Billon, N., Monteiro, M. & Dani, C., 2008. Developmental origin of adipocytes: new insights into a pending question. *Biology of the Cell*, 100(10), p.563.

Borchert, G.M., Lanier, W. & Davidson, B.L., 2006. RNA polymerase III transcribes human microRNAs. *Nat Struct Mol Biol*, 13(12), pp.1097-1101.

Bournat, J.C. & Brown, C.W., 2010. Mitochondrial dysfunction in obesity. *Current Opinion in Endocrinology, Diabetes, and Obesity*, 17(5), pp.446-452.

Buck, M.J. & Lieb, J.D., 2006. A chromatin-mediated mechanism for specification of conditional transcription factor targets. *Nat Genet*, 38(12), pp.1446-1451.

Cai, X., Hagedorn, C.H. & Cullen, B.R., 2004. Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA (New York, N.Y.)*, 10(12), pp.1957-1966.

Cali, J.J. et al., 2008. Bioluminescent assays for ADMET. *Expert Opinion on Drug Metabolism & Toxicology*, 4(1), pp.103-120.

Camargo, F.D. et al., 2003. Single hematopoietic stem cells generate skeletal muscle through myeloid intermediates. *Nature Medicine*, 9(12), pp.1520-1527.

Cao, H. et al., 2008. Identification of a Lipokine, a Lipid Hormone Linking Adipose Tissue to Systemic Metabolism. *Cell*, 134(6), pp.933-944.

Carninci, P., 2010. RNA Dust: Where are the Genes? *DNA Res*, p.dsq006.

Caygill, E.E. & Johnston, L.A., 2008. Temporal regulation of metamorphic processes in Drosophila by the let-7 and miR-125 heterochronic microRNAs. *Current Biology: CB*, 18(13), pp.943-950.

Ceccarini, C. & Eagle, H., 1971. pH as a determinant of cellular growth and contact inhibition. *Proceedings of the National Academy of Sciences of the United States of America*, 68(1), pp.229-233.

Chalfie, M., 1995. Green fluorescent protein. *Photochemistry and Photobiology*, 62(4), pp.651-656.

Chalfie, M., Horvitz, H.R. & Sulston, J.E., 1981. Mutations that lead to reiterations in the cell lineages of C. elegans. *Cell*, 24(1), pp.59-69.

Chatterjee, S. & Grosshans, H., 2009. Active turnover modulates mature microRNA activity in Caenorhabditis elegans. *Nature*, 461(7263), pp.546-549.

Chatterjee, S. & Pal, J., 2009. Role of 5'- and 3'-untranslated regions of mRNAs in human diseases. *Biology of the Cell*, 101(5), p.251.

Chen, C. et al., 2004. MicroRNAs modulate hematopoietic lineage differentiation. *Science (New York, N.Y.)*, 303(5654), pp.83-86.

Chen, L. et al., 2010. microRNA and aging: A novel modulator in regulating the aging network. *Ageing Research Reviews*. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20708718 [Accessed September 3, 2010].

Chendrimada, T.P. et al., 2007. MicroRNA silencing through RISC recruitment of eIF6. *Nature*, 447(7146), pp.823-828.

Cheng, C. & Li, L.M., Inferring MicroRNA Activities by Combining Gene Expression with MicroRNA Target Prediction. *PLoS ONE*, 3(4).

Chi, S.W. et al., 2009. Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature*, 460(7254), pp.479-486.

Chowdhury, E.H. et al., 2004. Dramatic effect of Mg2+ on transfecting mammalian cells by DNA/calcium phosphate precipitates. *Analytical Biochemistry*, 328(1), pp.96-97.

Cinti, S., 2006. The role of brown adipose tissue in human obesity. *Nutrition, Metabolism, and Cardiovascular Diseases: NMCD*, 16(8), pp.569-574.

Claverie, J., 2005. Fewer genes, more noncoding RNA. *Science (New York, N.Y.)*, 309(5740), pp.1529-1530.

Cohade, C. et al., 2003. Uptake in supraclavicular area fat ("USA-Fat"): description on 18F-FDG PET/CT. *Journal of Nuclear Medicine: Official Publication, Society of Nuclear Medicine*, 44(2), pp.170-176.

Cohen, B. et al., 2009. Ferritin nanoparticles as magnetic resonance reporter gene. *Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnology*, 1(2), pp.181-188.

Corcoran, D.L. et al., 2009. Features of mammalian microRNA promoters emerge from polymerase II chromatin immunoprecipitation data. *PloS One*, 4(4), p.e5279.

Cui, Q. et al., 2007. MicroRNAs preferentially target the genes with high transcriptional regulation complexity. *Biochemical and Biophysical*

Research Communications, 352(3), pp.733-738.

Damen, M. et al., 2010. Delivery of DNA and siRNA by novel gemini-like amphiphilic peptides. *Journal of Controlled Release: Official Journal of the Controlled Release Society*, 145(1), pp.33-39.

De, A., Loening, A.M. & Gambhir, S.S., 2007. An improved bioluminescence resonance energy transfer strategy for imaging intracellular events in single cells and living subjects. *Cancer Research*, 67(15), pp.7175-7183.

Doench, J.G., Petersen, C.P. & Sharp, P.A., 2003. siRNAs can function as miRNAs. *Genes Dev*, 17, pp.438-442.

Doench, J.G. & Sharp, P.A., 2004. Specificity of microRNA target selection in translational repression. *Genes & Development*, 18(5), pp.504-511.

Du, T. & Zamore, P.D., 2007. Beginning to understand microRNA function. *Cell Research*, 17(8), pp.661-663.

Easow, G., Teleman, A.A. & Cohen, S.M., 2007. Isolation of microRNA targets by miRNP immunopurification. *RNA (New York, N.Y.)*, 13(8), pp.1198-1204.

Elabd, C. et al., 2009. Human multipotent adipose-derived stem cells differentiate into functional brown adipocytes. *Stem Cells (Dayton, Ohio)*, 27(11), pp.2753-2760.

Elbashir, S.M., Lendeckel, W. & Tuschl, T., 2001. RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes & Development*, 15(2), pp.188-200.

Enright, A.J. et al., 2003. MicroRNA targets in Drosophila. *Genome Biology*, 5(1), p.R1.

Esau, C. et al., 2004. MicroRNA-143 regulates adipocyte differentiation. *The Journal of Biological Chemistry*, 279(50), pp.52361-52365.

Esslinger, S. & Förstemann, K., 2009. MicroRNAs repress mainly through mRNA decay. *Angewandte Chemie (International Ed. in English)*, 48(5), pp.853-855.

Eulalio, A., Behm-Ansmant, I. & Izaurralde, E., 2007. P bodies: at the crossroads of post-transcriptional pathways. *Nature Reviews. Molecular Cell Biology*, 8(1), pp.9-22.

Eulalio, A., Behm-Ansmant, I., Schweizer, D. et al., 2007. P-Body Formation Is a Consequence, Not the Cause, of RNA-Mediated Gene Silencing. *Molecular and Cellular Biology*, 27(11), pp.3970-3981.

Eulalio, A., Huntzinger, E. et al., 2009. Deadenylation is a widespread effect of miRNA regulation. *RNA (New York, N.Y.)*, 15(1), pp.21-32.

Eulalio, A., Tritschler, F. & Izaurralde, E., 2009. The GW182 protein family in animal cells: new insights into domains required for miRNA-mediated gene silencing. *RNA (New York, N.Y.)*, 15(8), pp.1433-1442.

Ewert, K. et al., 2005. Lipoplex structures and their distinct cellular pathways. *Advances in Genetics*, 53, pp.119-155.

Ewert, K. et al., 2004. Cationic lipid-DNA complexes for gene therapy: understanding the relationship between complex structure and gene delivery pathways at the molecular level. *Current Medicinal Chemistry*, 11(2), pp.133-149.

Ewert, K.K. et al., 2008. Non-viral gene delivery with cationic liposome-DNA complexes. *Methods in Molecular Biology (Clifton, N.J.)*, 433, pp.159-175.

Fan, F. & Wood, K.V., 2007. Bioluminescent assays for high-throughput screening. *Assay and Drug Development Technologies*, 5(1), pp.127-136.

Farmer, S.R., 2006. Transcriptional control of adipocyte formation. *Cell Metabolism*, 4(4), pp.263-273.

Fazi, F. & Nervi, C., 2008. MicroRNA: basic mechanisms and transcriptional regulatory networks for cell fate determination. *Cardiovascular Research*, 79(4), pp.553-561.

Feingold, K.R. et al., 1998. Paraoxonase activity in the serum and hepatic mRNA levels decrease during the acute phase response. *Atherosclerosis*, 139(2), pp.307-315.

Fernández-Riejos, P. et al., 2010. Role of leptin in the activation of immune cells. *Mediators of Inflammation*, 2010, p.568343.

de Ferranti, S. & Mozaffarian, D., 2008. The Perfect Storm: Obesity, Adipocyte Dysfunction, and Metabolic Consequences. *Clin Chem*, 54(6), pp.945-955.

Förstemann, K. et al., 2007. Drosophila microRNAs are sorted into functionally distinct argonaute complexes after production by dicer-1. *Cell*, 130(2), pp.287-297.

Friedländer, M.R. et al., 2008. Discovering microRNAs from deep sequencing data using miRDeep. *Nature Biotechnology*, 26(4), pp.407-415.

Friedman, R.C. et al., 2009. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Research*, 19(1), pp.92-105.

Furuhashi, M. & Hotamisligil, G.S., 2008. Fatty acid-binding proteins: role in metabolic diseases and potential as drug targets. *Nature Reviews. Drug Discovery*, 7(6), pp.489-503.

Gaidatzis, D. et al., 2007. Inference of miRNA targets using evolutionary conservation and pathway analysis. *BMC Bioinformatics*, 8, p.69.

Garzon, R., Marcucci, G. & Croce, C.M., 2010. Targeting microRNAs in cancer: rationale, strategies and challenges. *Nat Rev Drug Discov*, 9(10), pp.775-789.

Gebeshuber, C.A., Zatloukal, K. & Martinez, J., 2009. miR-29a suppresses tristetraprolin, which is a regulator of epithelial polarity and metastasis. *EMBO Rep*, 10(4), pp.400-405.

Gerasimova, T.I., Byrd, K. & Corces, V.G., 2000. A Chromatin Insulator Determines the Nuclear Localization of DNA. *Molecular Cell*, 6(5), pp.1025-1035.

Gesta, S., Tseng, Y. & Kahn, C.R., 2007. Developmental origin of fat: tracking obesity to its source. *Cell*, 131(2), pp.242-256.

Gilad, A.A. et al., 2008. MR REPORTER GENES. *Journal of nuclear medicine : official publication, Society of Nuclear Medicine*, 49(12), pp.1905-1908.

Gingeras, T.R., 2007. Origin of phenotypes: genes and transcripts. *Genome Research*, 17(6), pp.682-690.

Giraldez, A.J. et al., 2005. MicroRNAs Regulate Brain Morphogenesis in Zebrafish. *Science*, 308(5723), pp.833-838.

Gnacińska, M. et al., 2009. Role of adipokines in complications related to obesity: a review. *Advances in Medical Sciences*, 54(2), pp.150-157.

Gonzalez, S., Pisano, D.G. & Serrano, M., 2008. Mechanistic principles of chromatin remodeling guided by siRNAs and miRNAs. *Cell Cycle (Georgetown, Tex.)*, 7(16), pp.2601-2608.

Gregor, M.F. et al., 2009. Endoplasmic reticulum stress is reduced in tissues of obese subjects after weight loss. *Diabetes*, 58(3), pp.693-700.

Griffiths-Jones, S. et al., 2008. miRBase: tools for microRNA genomics. *Nucleic Acids Research*, 36(Database issue), pp.D154-158.

Grimson, A. et al., 2007. MicroRNA Targeting Specificity in Mammals: Determinants beyond Seed Pairing. , 27(1), pp.91-105.

Grishok, A. et al., 2001. Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing. *Cell*, 106(1), pp.23-34.

Guhaniyogi, J. & Brewer, G., 2001. Regulation of mRNA stability in mammalian cells. *Gene*, 265(1-2), pp.11-23.

Hackenberg, M. et al., 2009. miRanalyzer: a microRNA detection and analysis tool for next-generation sequencing experiments. , 37(Web Server issue), pp.W68-W76.

Hackl, H. et al., 2004. Analysis of DNA microarray data. *Current Topics in Medicinal Chemistry*, 4(13), pp.1357-1370.

Hamdy, O., Porramatikul, S. & Al-Ozairi, E., 2006. Metabolic obesity: the paradox between visceral and subcutaneous fat. *Current Diabetes Reviews*, 2(4), pp.367-373.

Hamrick, M.W. et al., 2010. The Adipokine Leptin Increases Skeletal Muscle Mass and Significantly Alters Skeletal Muscle miRNA Expression Profile in Aged Mice. *Biochemical and Biophysical Research Communications*. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20800581 [Accessed September 3, 2010].

Han, J., Kim, D. & Morris, K.V., 2007. Promoter-associated RNA is required for RNA-directed transcriptional gene silencing in human cells. *Proceedings of the National Academy of Sciences of the United States of America*, 104(30), pp.12422-12427.

Han, J. et al., 2006. Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. *Cell*, 125(5), pp.887-901.

Han, J. et al., 2004. The Drosha-DGCR8 complex in primary microRNA processing. *Genes & Development*, 18(24), pp.3016-3027.

Hany, T.F. et al., 2002. Brown adipose tissue: a factor to consider in symmetrical tracer uptake in the neck and upper chest region. *European Journal of Nuclear Medicine and Molecular Imaging*, 29(10), pp.1393-1398.

Hawkins, P.G. & Morris, K.V., 2008. RNA and transcriptional modulation of gene expression. *Cell Cycle (Georgetown, Tex.)*, 7(5), pp.602-607.

He, L. & Hannon, G.J., 2004. MicroRNAs: small RNAs with a big role in gene regulation. *Nat Rev Genet*, 5, pp.522-531.

Hirsch, J. & Batchelor, B., 1976. Adipose tissue cellularity in human obesity. *Clinics in Endocrinology and Metabolism*, 5(2), pp.299-311.

Öhman, M., 2007. A-to-I editing challenger or ally to the microRNA process. *Biochimie*, 89(10), pp.1171-1176.

Hobert, O., 2008. Gene Regulation by Transcription Factors and MicroRNAs. *Science*, 319(5871), pp.1785-1786.

Hoshino, H., Nakajima, Y. & Ohmiya, Y., 2007. Luciferase-YFP fusion tag with enhanced emission for single-cell luminescence imaging. *Nature Methods*, 4(8), pp.637-639.

Hotamisligil, G.S., 2010. Endoplasmic reticulum stress and the inflammatory basis of metabolic disease. *Cell*, 140(6), pp.900-917.

Huang, J.C. et al., 2007. Using expression profiling data to identify human microRNA targets. *Nature Methods*, 4(12), pp.1045-1049.

Huang, Y. et al., 2010. The changes in brain metabolism in people with activated brown adipose tissue: A PET study. *NeuroImage*. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20682349 [Accessed August 21, 2010].

Hummasti, S. & Hotamisligil, G.S., 2010. Endoplasmic reticulum stress and inflammation in obesity and diabetes. *Circulation Research*, 107(5), pp.579-591.

Hutvágner, G. et al., 2001. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science (New York, N.Y.)*, 293(5531), pp.834-838.

Hwang, H., Wentzel, E.A. & Mendell, J.T., 2007. A hexanucleotide element directs microRNA nuclear import. *Science (New York, N.Y.)*, 315(5808), pp.97-100.

Hwang, H. & Mendell, J.T., 2006. MicroRNAs in cell proliferation, cell death, and tumorigenesis. *Br J Cancer*, 94(6), pp.776-780.

Inui, M., Martello, G. & Piccolo, S., 2010. MicroRNA control of signal transduction. *Nat Rev Mol Cell Biol*, [Epub ahead of print]. Available at: http://genome.tugraz.at/Intranet/Papers/20216554.pdf.

Iwasaki, S. & Tomari, Y., 2009. Argonaute-mediated translational repression (and activation). *Fly*, 3(3), pp.204-206.

Jackson, A.L. et al., 2006. Widespread siRNA "off-target" transcript silencing mediated by seed region sequence complementarity. *RNA (New York, N.Y.)*, 12(7), pp.1179-1187.

James, V. et al., 2010. LIM-domain proteins, LIMD1, Ajuba, and WTIP are required for microRNA-mediated gene silencing. *Proceedings of the National Academy of Sciences*, 107(28), pp.12499 -12504.

Janowski, B.A. et al., 2007. Activating gene expression in mammalian cells with promoter-targeted duplex RNAs. *Nature Chemical Biology*, 3(3), pp.166-173.

Jastroch, M. et al., 2010. Mitochondrial proton and electron leaks. *Essays in Biochemistry*, 47, pp.53-67.

Jennewein, C. et al., 2010. MicroRNA-27b contributes to lipopolysaccharide-mediated peroxisome proliferator-activated receptor gamma (PPARgamma) mRNA destabilization. *The Journal of Biological Chemistry*, 285(16), pp.11846-11853.

Ji, Z. et al., 2009. Progressive lengthening of 3' untranslated regions of mRNAs by alternative polyadenylation during mouse embryonic development. , 106(17), pp.7028-7033.

Ji, Z. & Tian, B., Reprogramming of 3' Untranslated Regions of mRNAs by Alternative Polyadenylation in Generation of Pluripotent Stem Cells from Different Cell Types. , 4(12).

Johnston, R.J. et al., 2005. MicroRNAs acting in a double-negative feedback loop to control a neuronal cell fate decision. *Proceedings of the National Academy of Sciences of the United States of America*, 102(35), pp.12449-12454.

Jones-Rhoades, M.W., Bartel, D.P. & Bartel, B., 2006. MicroRNAS and their regulatory roles in plants. *Annual Review of Plant Biology*, 57, pp.19-53.

Jordan, M. & Wurm, F., 2004. Transfection of adherent and suspended cells by calcium phosphate. *Methods*, 33(2), pp.136-143.

Juven-Gershon, T. et al., 2008. The RNA polymerase II core promoter - the gateway to transcription. *Current Opinion in Cell Biology*, 20(3), pp.253-259.

Kajimura, S. et al., 2009. Initiation of myoblast to brown fat switch by a PRDM16–C/EBP-[bgr] transcriptional complex. *Nature*, 460(7259), pp.1154-1158.

Karbiener, M. et al., 2009. microRNA miR-27b impairs human adipocyte differentiation and targets PPARgamma. *Biochemical and Biophysical Research Communications*, 390(2), pp.247-251.

Kaur, H. et al., 2006. Thermodynamic, counterion, and hydration effects for the incorporation of locked nucleic acid nucleotides into DNA duplexes. *Biochemistry*, 45(23), pp.7347-7355.

Kawahara, Y. et al., 2007. Redirection of silencing targets by adenosine-to-inosine editing of miRNAs. *Science*, 315, pp.1137-1140.

Kawahara, Y. et al., 2007. RNA editing of the microRNA-151 precursor blocks cleavage by the Dicer-TRBP complex. *EMBO Reports*, 8(8), pp.763-769.

Kedde, M. & Agami, R., 2008. Interplay between microRNAs and RNA-binding proteins determines developmental processes. *Cell Cycle (Georgetown, Tex.)*, 7(7), pp.899-903.

Kertesz, M. et al., 2007. The role of site accessibility in microRNA target recognition. *Nature Genetics*, 39(10), pp.1278-1284.

Ketting, R.F. et al., 2001. Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. *Genes & Development*, 15(20), pp.2654-2659.

Khachane, A.N. & Harrison, P.M., Assessing the genomic evidence for conserved transcribed pseudogenes under selection. *BMC Genomics*, 10, pp.435-435.

Khateeb, J. et al., 2010. Paraoxonase 1 (PON1) expression in hepatocytes is upregulated by pomegranate polyphenols: A role for PPAR-? pathway. *Atherosclerosis*, 208(1), pp.119-125.

Kim, D.H. et al., 2008. MicroRNA-directed transcriptional gene silencing in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America*, 105(42), pp.16230-16235.

Kim, D.H. et al., 2006. Argonaute-1 directs siRNA-mediated transcriptional gene silencing in human cells. *Nature Structural & Molecular Biology*, 13(9), pp.793-797.

Kim, H. et al., 2010. Enhanced siRNA delivery using cationic liposomes with new polyarginine-conjugated PEG-lipid. *International Journal of Pharmaceutics*, 392(1-2), pp.141-147.

Kim, J. et al., 2007. Obesity-associated improvements in metabolic profile through expansion of adipose tissue. *The Journal of Clinical Investigation*, 117(9), pp.2621-2637.

Kim, V.N., 2004. MicroRNA precursors in motion: exportin-5 mediates their nuclear export. *Trends in Cell Biology*, 14(4), pp.156-159.

Kim, V.N., Han, J. & Siomi, M.C., 2009. Biogenesis of small RNAs in animals. *Nature Reviews. Molecular Cell Biology*, 10(2), pp.126-139.

Kim, V.N. & Nam, J., 2006. Genomics of microRNA. *Trends in Genetics: TIG*, 22(3), pp.165-173.

Kim, Y. & Kim, V.N., 2007. Processing of intronic microRNAs. *The EMBO Journal*, 26(3), pp.775-783.

Kiriakidou, M. et al., 2004. A combined computational-experimental approach predicts human microRNA targets. *Genes & Development*, 18(10), pp.1165-1178.

Kiriakidou, M. et al., 2007. An mRNA m7G cap binding-like motif within human Ago2 represses translation. *Cell*, 129(6), pp.1141-1151.

Kissebah, A.H. & Krakower, G.R., 1994. Regional adiposity and morbidity. *Physiological Reviews*, 74(4), pp.761-811.

Kloosterman, W.P. et al., 2004. Substrate requirements for let-7 function in the developing zebrafish embryo. *Nucleic Acids Research*, 32(21), pp.6284-6291.

Klöting, N. et al., 2009. MicroRNA expression in human omental and subcutaneous adipose tissue. *PloS One*, 4(3), p.e4699.

Kolachala, V.L. et al., 2010. Adenosine 2B receptor expression is post-transcriptionally regulated by microRNA. *The Journal of Biological Chemistry*, 285(24), pp.18184-18190.

Kong, W. et al., 2009. Strategies for profiling MicroRNA expression. *Journal of Cellular Physiology*, 218(1), pp.22-25.

Kreiss, P. et al., 1999. Plasmid DNA size does not affect the physicochemical properties of lipoplexes but modulates gene transfer efficiency. *Nucleic Acids Research*, 27(19), pp.3792-3798.

Krek, A. et al., 2005. Combinatorial microRNA target predictions. *Nature Genetics*, 37(5), pp.495-500.

Krichevsky, A.M. et al., 2003. A microRNA array reveals extensive regulation of microRNAs during brain development. *RNA (New York, N.Y.)*, 9(10), pp.1274-1281.

Kuk, J.L. et al., 2009. Age-related changes in total and regional fat distribution. *Ageing Research Reviews*, 8(4), pp.339-348.

Kumar, R. et al., 1998. The first analogues of LNA (locked nucleic acids): phosphorothioate-LNA and 2'-thio-LNA. *Bioorganic & Medicinal Chemistry Letters*, 8(16), pp.2219-2222.

Lagos-Quintana, M. et al., 2001. Identification of novel genes coding for small expressed RNAs. *Science (New York, N.Y.)*, 294(5543), pp.853-858.

Lagos-Quintana, M. et al., 2002. Identification of tissue-specific microRNAs from mouse. *Current Biology: CB*, 12(9), pp.735-739.

Lai, E.C. et al., 2003. Computational identification of Drosophila microRNA genes. *Genome Biology*, 4(7), p.R42.

Lau, D.C.W. et al., 2005. Adipokines: molecular links between obesity and atheroslcerosis. *Am J Physiol Heart Circ Physiol*, 288(5), pp.H2031-2041.

Lau, N.C. et al., 2001. An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. *Science (New York, N.Y.)*,

294(5543), pp.858-862.

Lee, D. et al., 2009. Getting the message across: mechanisms of physiological cross talk by adipose tissue. *American Journal of Physiology. Endocrinology and Metabolism*, 296(6), pp.E1210-1229.

Lee, I. et al., 2009. New class of microRNA targets containing simultaneous 5'-UTR and 3'-UTR interaction sites. , 19(7), pp.1175-1183.

Lee, R.C., Feinbaum, R.L. & Ambros, V., 1993. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell*, 75(5), pp.843-854.

Lee, Y.S. & Dutta, A., 2009. MicroRNAs in cancer. *Annual Review of Pathology*, 4, pp.199-227.

Lee, Y. et al., 2003. The nuclear RNase III Drosha initiates microRNA processing. *Nature*, 425(6956), pp.415-419.

Lee, Y. et al., 2006. The role of PACT in the RNA silencing pathway. *The EMBO Journal*, 25(3), pp.522-532.

Lee, Y. et al., 2002. MicroRNA maturation: stepwise processing and subcellular localization. *The EMBO Journal*, 21(17), pp.4663-4670.

Lee, Y. et al., 2004. MicroRNA genes are transcribed by RNA polymerase II. *The EMBO Journal*, 23(20), pp.4051-4060.

Lehrke, M. & Lazar, M.A., 2005. The many faces of PPARgamma. *Cell*, 123(6), pp.993-999.

Levy, S. et al., 2007. The Diploid Genome Sequence of an Individual Human. *PLoS Biology*, 5(10).

Lewis, B.P. et al., 2003a. Prediction of mammalian microRNA targets. *Cell*, 115(7), pp.787-798.

Lewis, B.P. et al., 2003b. Prediction of mammalian microRNA targets. *Cell*, 115(7), pp.787-798.

Li, G. & Widom, J., 2004. Nucleosomes facilitate their own invasion. *Nat Struct Mol Biol*, 11(8), pp.763-769.

Li, L. et al., 2010. Computational approaches for microRNA studies: a review. *Mammalian Genome: Official Journal of the International Mammalian Genome Society*, 21(1-2), pp.1-12.

Li, Q. et al., 2002. Locus control regions. *Blood*, 100(9), pp.3077-3086.

Li, T. et al., 2010. Identification of miR-130a, miR-27b and miR-210 as serum biomarkers for atherosclerosis obliterans. *Clinica Chimica Acta; International Journal of Clinical Chemistry*. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20888330 [Accessed October 10, 2010].

Licatalosi, D.D. et al., 2008. HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature*, 456(7221), pp.464-469.

Lijnen, H.R., 2005. Pleiotropic functions of plasminogen activator inhibitor-1. *Journal of Thrombosis and Haemostasis: JTH*, 3(1), pp.35-45.

Lim, L.P. et al., 2005. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature*, 433(7027), pp.769-773.

Lim, L.P. et al., 2003. The microRNAs of Caenorhabditis elegans. *Genes & Development*, 17(8), pp.991-1008.

Liu, H.S. et al., 1999. Is green fluorescent protein toxic to the living cells? *Biochemical and Biophysical Research Communications*, 260(3), pp.712-717.

Loening, A.M., Fenn, T.D. & Gambhir, S.S., 2007. Crystal structures of the luciferase and green fluorescent protein from Renilla reniformis. *Journal of Molecular Biology*, 374(4), pp.1017-1028.

Loening, A.M. et al., 2006. Consensus guided mutagenesis of Renilla luciferase yields enhanced stability and light output. *Protein Engineering, Design & Selection: PEDS*, 19(9), pp.391-400.

Lonez, C., Vandenbranden, M. & Ruysschaert, J., 2008. Cationic liposomal lipids: from gene carriers to cell signaling. *Progress in Lipid Research*, 47(5), pp.340-347.

Long, D. et al., 2007. Potent effect of target structure on microRNA function. *Nature Structural & Molecular Biology*, 14(4), pp.287-294.

Louis, N., Evelegh, C. & Graham, F.L., 1997. Cloning and sequencing of the cellular-viral junctions from the human adenovirus type 5 transformed 293 cell line. *Virology*, 233(2), pp.423-429.

Lund, E. et al., 2004. Nuclear export of microRNA precursors. *Science (New York, N.Y.)*, 303(5654), pp.95-98.

Luo, N., Liu, J., Chung, B.H. et al., 2010. Macrophage adiponectin expression improves insulin sensitivity and protects against inflammation and atherosclerosis. *Diabetes*, 59(4), pp.791-799.

Luo, N., Liu, J., Chung, B.H. et al., 2010. Macrophage Adiponectin Expression Improves Insulin Sensitivity and Protects Against Inflammation and Atherosclerosis. *Diabetes*, 59(4), pp.791 -799.

Lusis, A.J., 2000. Atherosclerosis. *Nature*, 407(6801), pp.233-241.

Ma, Z. et al., 2007. Quantitative analysis of cotransfection efficiencies in studies of ionotropic glutamate receptor complexes. *Journal of Neuroscience Research*, 85(1), pp.99-115.

MacRae, I.J. et al., 2008. In vitro reconstitution of the human RISC-loading complex. *Proceedings of the National Academy of Sciences of the United States of America*, 105(2), pp.512-517.

Macville, M. et al., 1999. Comprehensive and definitive molecular cytogenetic characterization of HeLa cells by spectral karyotyping. *Cancer Research*, 59(1), pp.141-150.

Majka, S.M. et al., 2010. De novo generation of white adipocytes from the myeloid lineage via mesenchymal intermediates is age, adipose depot, and gender specific. *Proceedings of the National Academy of Sciences of the United States of America*, 107(33), pp.14781-14786.

Majoros, W.H. & Ohler, U., Spatial preferences of microRNA targets in 3' untranslated regions. *BMC Genomics*, 8, pp.152-152.

Makeyev, E.V. et al., 2007. The MicroRNA miR-124 promotes neuronal differentiation by triggering brain-specific alternative pre-mRNA splicing. *Molecular Cell*, 27(3), pp.435-448.

Mangan, S. & Alon, U., 2003. Structure and function of the feed-forward loop network motif. *Proceedings of the National Academy of Sciences of the United States of America*, 100(21), pp.11980-11985.

van Marken Lichtenbelt, W.D. et al., 2009. Cold-activated brown adipose tissue in healthy men. *The New England Journal of Medicine*, 360(15), pp.1500-1508.

Marson, A. et al., 2008. Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell*, 134(3), pp.521-533.

Martinez, J. & Tuschl, T., 2004. RISC is a 5' phosphomonoester-producing RNA endonuclease. *Genes & Development*, 18(9), pp.975-980.

Marx, N. et al., 1998. Macrophages in human atheroma contain PPARgamma: differentiation-dependent peroxisomal proliferator-activated receptor gamma(PPARgamma) expression and reduction of MMP-9 activity through PPARgamma activation in mononuclear phagocytes in vitro. *The American Journal of Pathology*, 153(1), pp.17-23.

Mazumder, B., Seshadri, V. & Fox, P.L., 2003. Translational control by the 3'-UTR: the ends specify the means. *Trends in Biochemical Sciences*, 28(2), pp.91-98.

McKenna, N.J. & O'Malley, B.W., 2002. Minireview: nuclear receptor coactivators--an update. *Endocrinology*, 143(7), pp.2461-2465.

Mendes, N.D., Freitas, A.T. & Sagot, M., 2009. Current tools for the identification of miRNA genes and their targets. *Nucleic Acids Research*, 37(8), pp.2419-2433.

Millevoi, S. & Vagner, S., 2010. Molecular mechanisms of eukaryotic pre-mRNA 3′ end processing regulation. , 38(9), pp.2757-2774.

Monteys, A.M. et al., 2010. Structure and activity of putative intronic miRNA promoters. *RNA (New York, N.Y.)*, 16(3), pp.495-505.

Morin, R.D. et al., 2008. Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. *Genome Research*, 18(4), pp.610-621.

Morlando, M. et al., 2008. Primary microRNA transcripts are processed co-transcriptionally. *Nat Struct Mol Biol*, 15(9), pp.902-909.

Morris, K.V. et al., 2008. Bidirectional transcription directs both transcriptional gene activation and suppression in human cells. *PLoS Genetics*, 4(11), p.e1000258.

Nakatsu, T. et al., 2006. Structural basis for the spectral difference in luciferase bioluminescence. *Nature*, 440(7082), pp.372-376.

Nedergaard, J., Bengtsson, T. & Cannon, B., 2007. Unexpected evidence for active brown adipose tissue in adult humans. *American Journal of Physiology. Endocrinology and Metabolism*, 293(2), pp.E444-452.

Nicolakakis, N. & Hamel, E., 2010. The Nuclear Receptor PPARgamma as a Therapeutic Target for Cerebrovascular and Brain Dysfunction in Alzheimer's Disease. *Frontiers in Aging Neuroscience*, 2. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20725514 [Accessed September 24, 2010].

Nielsen, C.B. et al., 2007. Determinants of targeting by endogenous and exogenous microRNAs and siRNAs. *RNA*, 13, pp.1894-1910.

van Nimwegen, E., 2003. Scaling laws in the functional content of genomes. *Trends in Genetics: TIG*, 19(9), pp.479-484.

Nishikura, K., 2010. Functions and regulation of RNA editing by ADAR deaminases. *Annual Review of Biochemistry*, 79, pp.321-349.

Obernosterer, G. et al., 2006. Post-transcriptional regulation of microRNA expression. *RNA*. Available at: http://genome.tugraz.at/Intranet/Papers/16738409.pdf.

O'Connell, R.M. et al., 2010. Physiological and pathological roles for microRNAs in the immune system. *Nature Reviews. Immunology*, 10(2), pp.111-122.

Ogawa, R. et al., 2010. Adipocyte-derived microvesicles contain RNA that is transported into macrophages and might be secreted into blood circulation. *Biochemical and Biophysical Research Communications*, 398(4), pp.723-729.

Ohler, U. et al., 2004. Patterns of flanking sequence conservation and a characteristic upstream motif for microRNA gene identification. *RNA (New York, N.Y.)*, 10(9), pp.1309-1322.

Okamura, K., Chung, W. & Lai, E.C., 2008. The long and short of inverted repeat genes in animals: microRNAs, mirtrons and hairpin RNAs. *Cell Cycle (Georgetown, Tex.)*, 7(18), pp.2840-2845.

Okamura, K. et al., 2007. The mirtron pathway generates microRNA-class regulatory RNAs in Drosophila. *Cell*, 130(1), pp.89-100.

Ozsolak, F. et al., 2008. Chromatin structure analyses identify miRNA promoters. *Genes & Development*, 22(22), pp.3172-3183.

Ozsolak, F. et al., 2007. High-throughput mapping of the chromatin structure of human promoters. *Nature Biotechnology*, 25(2), pp.244-248.

Pantano, L., Estivill, X. & Martí, E., 2010. SeqBuster, a bioinformatic tool for the processing and analysis of small RNAs datasets, reveals ubiquitous miRNA modifications in human embryonic cells. *Nucleic Acids Research*, 38(5), p.e34.

Park, S. et al., 2009. miR-29 miRNAs activate p53 by targeting p85 alpha and CDC42. *Nature Structural & Molecular Biology*, 16(1), pp.23-29.

Parker, R. & Sheth, U., 2007. P bodies and the control of mRNA translation and degradation. *Molecular Cell*, 25(5), pp.635-646.

Pasquinelli, A.E. et al., 2000. Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature*, 408(6808), pp.86-89.

Paulmurugan, R. & Gambhir, S.S., 2003. Monitoring protein-protein interactions using split synthetic renilla luciferase protein-fragment-assisted complementation. *Analytical Chemistry*, 75(7), pp.1584-1589.

Pelton, P.D. et al., 1999. PPARgamma activation induces the expression of the adipocyte fatty acid binding protein gene in human monocytes. *Biochemical and Biophysical Research Communications*, 261(2), pp.456-458.

Pénicaud, L., 2010. The neural feedback loop between the brain and adipose tissues. *Endocrine Development*, 19, pp.84-92.

Pesole, G. et al., 2001. Structural and functional features of eukaryotic mRNA untranslated regions. *Gene*, 276(1-2), pp.73-81.

Peters, L. & Meister, G., 2007. Argonaute proteins: mediators of RNA silencing. *Molecular Cell*, 26(5), pp.611-623.

Pfannenberg, C. et al., 2010. Impact of age on the relationships of brown adipose tissue with sex and adiposity in humans. *Diabetes*, 59(7), pp.1789-1793.

Philpott, N. et al., 1996. The use of 7-amino actinomycin D in identifying apoptosis: simplicity of use and broad spectrum of application compared with other techniques. *Blood*, 87(6), pp.2244-2251.

Pickering, B.M. & Willis, A.E., 2005. The implications of structured 5' untranslated regions on translation and disease. *Seminars in Cell & Developmental Biology*, 16(1), pp.39-47.

Pieler, R. et al., 2004. ArrayNorm: comprehensive normalization and analysis of microarray data. *Bioinformatics (Oxford, England)*, 20(12), pp.1971-1973.

Pinto-Gonzalez Howell, D. et al., 2003. Deoxyribonuclease II is a Lysosomal Barrier to Transfection. *Mol Ther*, 8(6), pp.957-963.

Place, R.F. et al., 2008. MicroRNA-373 induces expression of genes with complementary promoter sequences. *Proceedings of the National Academy of Sciences of the United States of America*, 105(5), pp.1608-1613.

Prasanth, K.V. & Spector, D.L., 2007. Eukaryotic regulatory RNAs: an answer to the 'genome complexity' conundrum. *Genes & Development*, 21(1), pp.11-42.

Puri, A. et al., 2009. Lipid-Based Nanoparticles as Pharmaceutical Drug Carriers: From Concepts to Clinic. *Critical reviews in therapeutic drug carrier systems*, 26(6), pp.523-580.

Quackenbush, J., 2002. Microarray data normalization and transformation. *Nature Genetics*, 32 Suppl, pp.496-501.

Rajewsky, N., 2006a. microRNA target predictions in animals. *Nature Genetics*, 38 Suppl, pp.S8-13.

Rajewsky, N., 2006b. microRNA target predictions in animals. *Nature Genetics*, 38 Suppl, pp.S8-13.

Ribeiro, M.O. et al., 2010. Expression of uncoupling protein 1 in mouse brown adipose tissue is thyroid hormone receptor-beta isoform specific and required for adaptive thermogenesis. *Endocrinology*, 151(1), pp.432-440.

RIKEN Genome Exploration Research Group and Genome Science Group (Genome Network Project Core Group) and the FANTOM

Consortium et al., 2005. Antisense Transcription in the Mammalian Transcriptome. *Science*, 309(5740), pp.1564-1566.

Ro, S. et al., 2007. Tissue-dependent paired expression of miRNAs. *Nucleic Acids Research*, 35(17), pp.5944-5953.

Robb, G.B. et al., 2005. Specific and potent RNAi in the nucleus of human cells. *Nature Structural & Molecular Biology*, 12(2), pp.133-137.

Rodriguez, A. et al., 2005. The human adipose tissue is a source of multipotent stem cells. *Biochimie*, 87(1), pp.125-128.

Rodriguez, A. et al., 2004. Adipocyte differentiation of multipotent cells established from human adipose tissue. *Biochemical and Biophysical Research Communications*, 315(2), pp.255-263.

Rodriguez, A. et al., 2004. Identification of mammalian microRNA host genes and transcription units. *Genome Research*, 14(10A), pp.1902-1910.

Ronaghi, M. et al., 1996. Real-time DNA sequencing using detection of pyrophosphate release. *Analytical Biochemistry*, 242(1), pp.84-89.

Rong, J.X. et al., 2007. Adipose mitochondrial biogenesis is suppressed in db/db and high-fat diet-fed mice and improved by rosiglitazone. *Diabetes*, 56(7), pp.1751-1760.

Rosen, E.D. & MacDougald, O.A., 2006. Adipocyte differentiation from the inside out. *Nature Reviews. Molecular Cell Biology*, 7(12), pp.885-896.

Rousseau, C. et al., 2006. Brown fat in breast cancer patients: analysis of serial (18)F-FDG PET/CT scans. *European Journal of Nuclear Medicine and Molecular Imaging*, 33(7), pp.785-791.

Ruby, J.G. et al., 2006. Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in C. elegans. *Cell*, 127(6), pp.1193-1207.

Ruby, J.G., Jan, C.H. & Bartel, D.P., 2007. Intronic microRNA precursors that bypass Drosha processing. *Nature*, 448(7149), pp.83-86.

Rusinov, V. et al., 2005. MicroInspector: a web tool for detection of miRNA binding sites in an RNA sequence. *Nucleic Acids Research*, 33(Web Server issue), pp.W696-700.

Ruvkun, G., 2001. Molecular biology. Glimpses of a tiny RNA world. *Science (New York, N.Y.)*, 294(5543), pp.797-799.

Saetrom, P. et al., 2007. Distance constraints between microRNA target sites dictate efficacy and cooperativity. *Nucleic Acids Res*, 35, pp.2333-2342.

Saini, H.K., Enright, A.J. & Griffiths-Jones, S., 2008. Annotation of mammalian primary microRNAs. *BMC Genomics*, 9, p.564.

Saini, H.K., Griffiths-Jones, S. & Enright, A.J., 2007. Genomic analysis of human microRNA transcripts. *Proceedings of the National Academy of Sciences of the United States of America*, 104(45), pp.17719-17724.

Sakamoto, S. et al., 2009. The NF90-NF45 complex functions as a negative regulator in the microRNA processing pathway. *Molecular and Cellular Biology*, 29(13), pp.3754-3769.

Sandelin, A. et al., 2007. Mammalian RNA polymerase II core promoters: insights from genome-wide studies. *Nature Reviews Genetics*, 8(6), pp.424-436.

Sasaki, T. et al., 2003. Identification of eight members of the Argonaute family in the human genome small star, filled. *Genomics*, 82(3), pp.323-330.

Schmid, I. et al., 1992. Dead cell discrimination with 7-amino-actinomycin D in combination with dual color immunofluorescence in single laser flow cytometry. *Cytometry*, 13(2), pp.204-208.

Schones, D.E. et al., 2008. Dynamic regulation of nucleosome positioning in the human genome. *Cell*, 132(5), pp.887-898.

Schwartz, J.C. et al., 2008. Antisense transcripts are targets for activating small RNAs. *Nature Structural & Molecular Biology*, 15(8), pp.842-848.

Schwarz, D.S. et al., 2003. Asymmetry in the assembly of the RNAi enzyme complex. *Cell*, 115(2), pp.199-208.

Seale, P. et al., 2008. PRDM16 controls a brown fat/skeletal muscle switch. *Nature*, 454(7207), pp.961-967.

Selbach, M. et al., 2008. Widespread changes in protein synthesis induced by microRNAs. *Nature*, [Epub ahead of print]. Available at: http://genome.tugraz.at/Intranet/Papers/18668040.pdf.

Sempere, L.F. et al., 2004. Expression profiling of mammalian microRNAs uncovers a subset of brain-expressed microRNAs with possible roles in murine and human neuronal differentiation. *Genome Biology*, 5(3), pp.R13-R13.

Senn, J.J. et al., 2002. Interleukin-6 induces cellular insulin resistance in hepatocytes. *Diabetes*, 51(12), pp.3391-3399.

Sethupathy, P., Megraw, M. & Hatzigeorgiou, A.G., 2006. A guide through present computational approaches for the identification of

mammalian microRNA targets. *Nature Methods*, 3(11), pp.881-886.

Sharova, L.V. et al., 2009. Database for mRNA Half-Life of 19 977 Genes Obtained by DNA Microarray Analysis of Pluripotent and Differentiating Mouse Embryonic Stem Cells. *DNA Research*, 16(1), pp.45 -58.

Shaw, G. et al., 2002. Preferential transformation of human neuronal cells by human adenoviruses and the origin of HEK 293 cells. *The FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology*, 16(8), pp.869-871.

Shibusawa, N., Hashimoto, K. & Yamada, M., 2008. Thyrotropin-releasing hormone (TRH) in the cerebellum. *Cerebellum (London, England)*, 7(1), pp.84-95.

Shu, X. et al., 2009. Mammalian expression of infrared fluorescent proteins engineered from a bacterial phytochrome. *Science (New York, N.Y.)*, 324(5928), pp.804-807.

Siersbaek, R., Nielsen, R. & Mandrup, S., 2010. PPARgamma in adipocyte differentiation and metabolism--novel insights from genome-wide studies. *FEBS Letters*, 584(15), pp.3242-3249.

Smale, S.T., 2010a. Beta-galactosidase assay. *Cold Spring Harbor Protocols*, 2010(5), p.pdb.prot5423.

Smale, S.T., 2010b. Chloramphenicol acetyltransferase assay. *Cold Spring Harbor Protocols*, 2010(5), p.pdb.prot5422.

So, M. et al., 2006. Self-illuminating quantum dot conjugates for in vivo imaging. *Nature Biotechnology*, 24(3), pp.339-343.

Song Gao, J. et al., 2010. Atypical transcription of microRNA gene fragments. *Nucleic Acids Research*. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20097657 [Accessed March 25, 2010].

Spagnou, S., Miller, A.D. & Keller, M., 2004. Lipidic carriers of siRNA: differences in the formulation, cellular uptake, and delivery with plasmid DNA. *Biochemistry*, 43(42), pp.13348-13356.

Spilianakis, C.G. et al., 2005. Interchromosomal associations between alternatively expressed loci. *Nature*, 435(7042), pp.637-645.

Stanley, S. et al., 2010. Identification of neuronal subpopulations that project from hypothalamus to both liver and adipose tissue polysynaptically. *Proceedings of the National Academy of Sciences of the United States of America*, 107(15), pp.7024-7029.

Steiner, F.A. et al., 2007. Structural features of small RNA precursors determine Argonaute loading in Caenorhabditis elegans. *Nature Structural & Molecular Biology*, 14(10), pp.927-933.

Steitz, J.A. & Vasudevan, S., 2009. miRNPs: versatile regulators of gene expression in vertebrate cells. *Biochemical Society Transactions*, 37(Pt 5), pp.931-935.

Stofkova, A., 2009. Leptin and adiponectin: from energy and metabolic dysbalance to inflammation and autoimmunity. *Endocrine Regulations*, 43(4), pp.157-168.

Sturn, A., Quackenbush, J. & Trajanoski, Z., 2002. Genesis: cluster analysis of microarray data. *Bioinformatics (Oxford, England)*, 18(1), pp.207-208.

Suyama, M. et al., 2006. Identification and Analysis of Genes and Pseudogenes within Duplicated Regions in the Human and Mouse Genomes. *PLoS Comput Biol*, 2(6), p.e76.

Suzuki, H.I. et al., 2009. Modulation of microRNA processing by p53. *Nature*, 460(7254), pp.529-533.

Suzuki, K. & Kelleher, A.D., 2009. Transcriptional regulation by promoter targeted RNAs. *Current Topics in Medicinal Chemistry*, 9(12), pp.1079-1087.

Sykes, M.L. & Avery, V.M., 2009. A luciferase based viability assay for ATP detection in 384-well format for high throughput whole cell screening of Trypanosoma brucei brucei bloodstream form strain 427. *Parasites & Vectors*, 2(1), p.54.

Tchkonia, T. et al., 2005. Abundance of two human preadipocyte subtypes with distinct capacities for replication, adipogenesis, and apoptosis varies among fat depots. *American Journal of Physiology. Endocrinology and Metabolism*, 288(1), pp.E267-277.

Thadani, R. & Tammi, M.T., 2006. MicroTar: predicting microRNA targets from RNA duplexes. *BMC Bioinformatics*, 7 Suppl 5, p.S20.

Thomson, J.M. et al., 2006. Extensive post-transcriptional regulation of microRNAs and its implications for cancer. *Genes & Development*, 20(16), pp.2202-2207.

Tomari, Y., Du, T. & Zamore, P.D., 2007. Sorting of Drosophila small silencing RNAs. *Cell*, 130(2), pp.299-308.

Tsang, J., Zhu, J. & van Oudenaarden, A., 2007. MicroRNA-Mediated Feedback and Feedforward Loops Are Recurrent Network Motifs in Mammals. , 26(5), pp.753-767.

Tseng, W., Haselton, F.R. & Giorgio, T.D., 1997. Transfection by Cationic Liposomes Using Simultaneous Single Cell Measurements of Plasmid Delivery and Transgene Expression. *Journal of Biological Chemistry*, 272(41), pp.25641 -25647.

Tsien, R.Y., 1998. The green fluorescent protein. *Annual Review of Biochemistry*, 67, pp.509-544.

Turner, M.J. & Slack, F.J., 2009. Transcriptional control of microRNA expression in C. elegans: promoting better understanding. *RNA Biology*, 6(1), pp.49-53.

Ukropec, J., Anunciado, R.P., Ravussin, Y., Hulver, M.W. et al., 2006. UCP1-independent Thermogenesis in White Adipose Tissue of Cold-acclimated Ucp1-/- Mice. *Journal of Biological Chemistry*, 281(42), pp.31894 -31908.

Ukropec, J., Anunciado, R.V.P., Ravussin, Y. & Kozak, L.P., 2006. Leptin is required for uncoupling protein-1-independent thermogenesis during cold stress. *Endocrinology*, 147(5), pp.2468-2480.

Vastenhouw, N.L. et al., 2006. Gene expression: long-term gene silencing by RNAi. *Nature*, 442(7105), p.882.

Vasudevan, S. & Steitz, J.A., 2007. AU-rich-element-mediated upregulation of translation by FXR1 and Argonaute 2. *Cell*, 128(6), pp.1105-1118.

Vasudevan, S., Tong, Y. & Steitz, J.A., 2007. Switching from repression to activation: microRNAs can up-regulate translation. *Science (New York, N.Y.)*, 318(5858), pp.1931-1934.

Venisnik, K.M. et al., 2006. Bifunctional antibody-Renilla luciferase fusion protein for in vivo optical detection of tumors. *Protein Engineering, Design & Selection: PEDS*, 19(10), pp.453-460.

Venter, J.C. et al., 2001. The Sequence of the Human Genome. *Science*, 291(5507), pp.1304-1351.

Ventura, A. et al., 2008. Targeted deletion reveals essential and overlapping functions of the miR-17 through 92 family of miRNA clusters. *Cell*, 132(5), pp.875-886.

Vidal, H., 2001. Gene expression in visceral and subcutaneous adipose tissues. *Annals of Medicine*, 33(8), pp.547-555.

Viswakarma, N. et al., 2010. Coactivators in PPAR-Regulated Gene Expression. *PPAR Research*, 2010. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20814439 [Accessed September 24, 2010].

Vohl, M. et al., 2004. A survey of genes differentially expressed in subcutaneous and visceral adipose tissue in men. *Obesity Research*, 12(8), pp.1217-1222.

Wang, J., 2005. Finding primary targets of transcriptional regulators. *Cell Cycle (Georgetown, Tex.)*, 4(3), pp.356-358.

Wang, T. & Xu, Z., 2010. miR-27 promotes osteoblast differentiation by modulating Wnt signaling. *Biochemical and Biophysical Research Communications*. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20708603 [Accessed October 18, 2010].

Wang, X. et al., 2009. High-resolution human core-promoter prediction with CoreBoost_HM. *Genome Research*, 19(2), pp.266-275.

Wang, Y. et al., 2010. Nanoparticle-based delivery system for application of siRNA in vivo. *Current Drug Metabolism*, 11(2), pp.182-196.

Weber, M.J., 2005. New human and mouse microRNA genes found by homology search. *The FEBS Journal*, 272(1), pp.59-73.

Widder, E.A., 2010. Bioluminescence in the Ocean: Origins of Biological, Chemical, and Ecological Diversity. *Science*, 328(5979), pp.704-708.

Wienholds, E. et al., 2005. MicroRNA Expression in Zebrafish Embryonic Development. *Science*, 309(5732), pp.310-311.

Wienholds, E. & Plasterk, R.H., 2005. MicroRNA function in animal development. *FEBS Letters*, 579(26), pp.5911-5922.

Wiesen, J.L. & Tomasi, T.B., 2009. Dicer is regulated by cellular stresses and interferons. *Molecular Immunology*, 46(6), pp.1222-1228.

Willenbring, H. et al., 2004. Myelomonocytic cells are sufficient for therapeutic cell fusion in liver. *Nature Medicine*, 10(7), pp.744-748.

Wolff, J.A. & Budker, V., 2005. The mechanism of naked DNA uptake and expression. *Advances in Genetics*, 54, pp.3-20.

Woo, J. & von Arnim, A.G., Mutational optimization of the coelenterazine-dependent luciferase from Renilla. *Plant Methods*, 4, pp.23-23.

Woods, S.C., Gotoh, K. & Clegg, D.J., 2003. Gender differences in the control of energy homeostasis. *Experimental Biology and Medicine (Maywood, N.J.)*, 228(10), pp.1175-1180.

Xie, X. et al., 2005. Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals. *Nature*, 434(7031), pp.338-345.

Xu, P. et al., 2003. The Drosophila microRNA Mir-14 suppresses cell death and is required for normal fat metabolism. *Current Biology: CB*, 13(9), pp.790-795.

Yang, F., Moss, L.G. & Phillips, G.N., 1996. The molecular structure of green fluorescent protein. *Nature Biotechnology*, 14(10), pp.1246-1251.

Yang, L. et al., 2010. Defective hepatic autophagy in obesity promotes ER stress and causes insulin resistance. *Cell Metabolism*, 11(6),

pp.467-478.

Yeaman, S.J., 2004. Hormone-sensitive lipase--new roles for an old enzyme. *The Biochemical Journal*, 379(Pt 1), pp.11-22.

YI, R. et al., 2005. Overexpression of Exportin 5 enhances RNA interference mediated by short hairpin RNAs and microRNAs. *RNA*, 11(2), pp.220-226.

Yi, R. et al., 2003. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes & Development*, 17(24), pp.3011-3016.

Yu, S. & Reddy, J.K., 2007. Transcription coactivators for peroxisome proliferator-activated receptors. *Biochimica Et Biophysica Acta*, 1771(8), pp.936-951.

Zabner, J. et al., 1995. Cellular and molecular barriers to gene transfer by a cationic lipid. *The Journal of Biological Chemistry*, 270(32), pp.18997-19007.

Zeng, Y. & Cullen, B.R., 2005. Efficient processing of primary microRNA hairpins by Drosha requires flanking nonstructured RNA sequences. *The Journal of Biological Chemistry*, 280(30), pp.27595-27603.

Zhang, B. et al., 2007. microRNAs as oncogenes and tumor suppressors. *Developmental Biology*, 302(1), pp.1-12.

Zhang, C., 2008. MicroRNAs: role in cardiovascular biology and disease. *Clinical Science (London, England: 1979)*, 114(12), pp.699-706.

Zhang, H. et al., 2002. Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP. *The EMBO Journal*, 21(21), pp.5875-5885.

Zheng, D. & Gerstein, M.B., 2007. The ambiguous boundary between genes and pseudogenes: the dead rise up, or do they? *Trends in Genetics*, 23(5), pp.219-224.

Zhou, X. et al., 2007. Characterization and identification of microRNA core promoters in four model species. *PLoS Computational Biology*, 3(3), p.e37.

Zhu, E. et al., 2010. mirTools: microRNA profiling and discovery based on high-throughput sequencing. *Nucleic Acids Research*. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20478827 [Accessed May 24, 2010].

Ziello, J.E., Huang, Y. & Jovin, I.S., 2010. Cellular Endocytosis and Gene Delivery. , 16(5-6), pp.222-229.

Books and online resources

Stryer L, Berg JM, Tymoczko JL. Biochemie. Spektrum Akademischer Verlag 2003 (5. Auflage), ISBN 3-8274-1303-6
Campbell NA, Reece JB. Biologie. Spektrum Akademischer Verlag 2003 (6. Auflage), ISBN 3-8274-1352-4

Opriessnig Peter. Identification of microRNA-mRNA interaction networks targeting adipogenesis and obesity. Diploma thesis, IGB 2008.

MPP – Molecular Plant Physiology. University of Erlangen online resource 2006. www.biologie.uni-erlangen.de/mpp
WHO 2010. www.who.int/topics/obesity
Promega T. Technical resources www.promega.com
Promega. www.promega.com
Biontex. www.biontex.com
Fermentas. www.fermentas.com
Exiqon. www.exiqon.com

8. Appendix

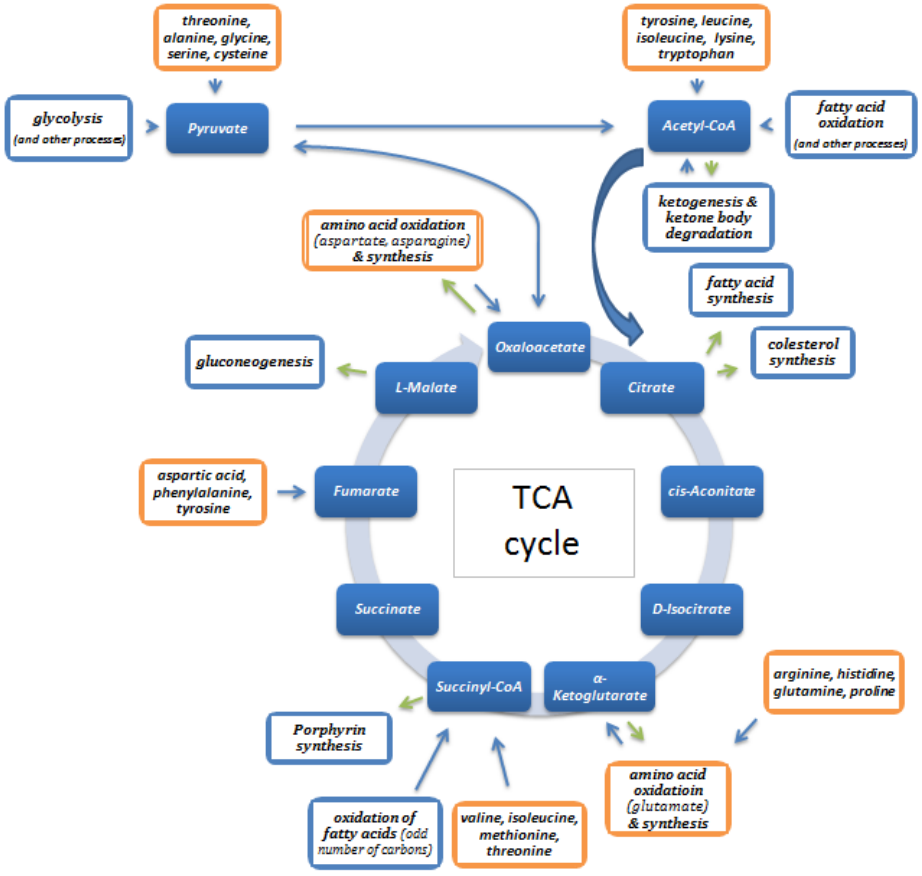


Fig. A1 Citric acid cycle. The potentially toxic nitrogen of amino acids is eliminated via transaminations, deamination, and urea formation. Carbon skeletons are generally conserved as carbohydrate (gluconeogenesis) or as fatty acid (fatty acid synthesis). In this respect amino acids are distinguishable into the categories *glucogenic* and *ketogenic*. Glucogenic amino acids are those that give rise to production of *pyruvate* or TCA cycle intermediates (like *α-ketoglutarate* and *oxaloacetate*), that are precursors to l-*malate* and thus *glucose* via gluconeogenesis. All amino acids except lysine and leucine are at least partly glucogenic. Lysine and leucine are solely ketogenic, giving rise only to *acetyl-CoA* or *acetoacetyl-CoA*.

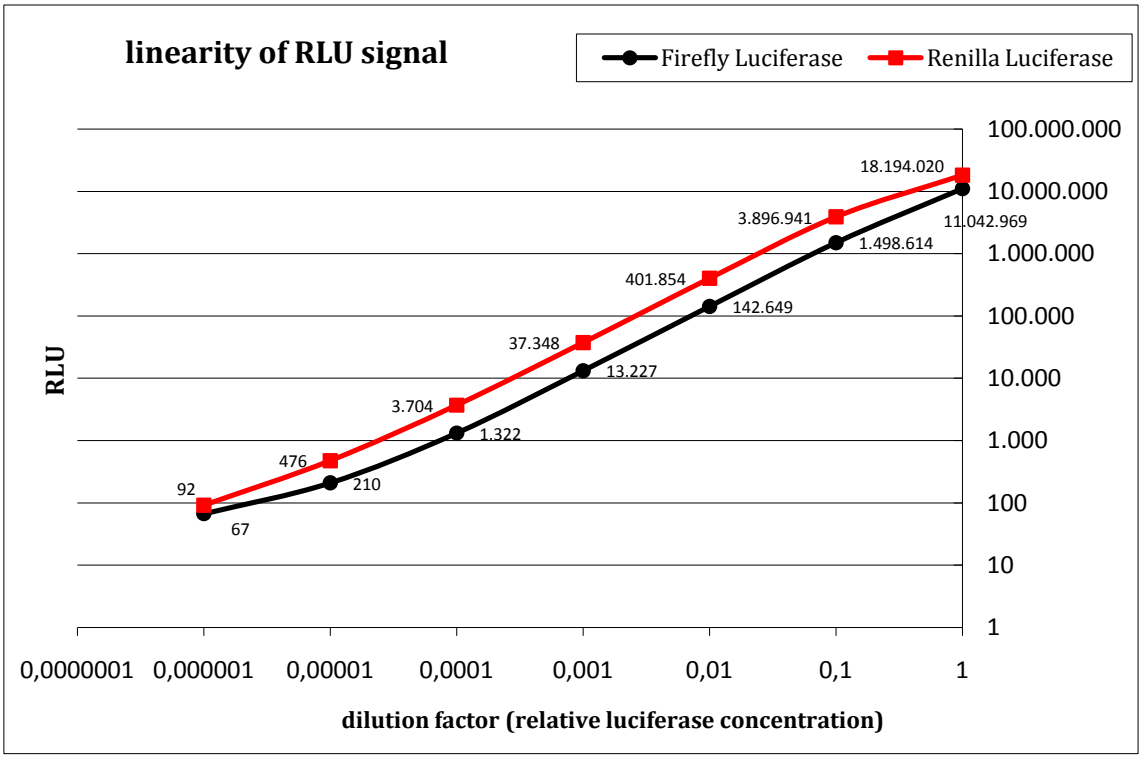


Fig. A2 **Linearity of RLU signal – full dynamic range**.

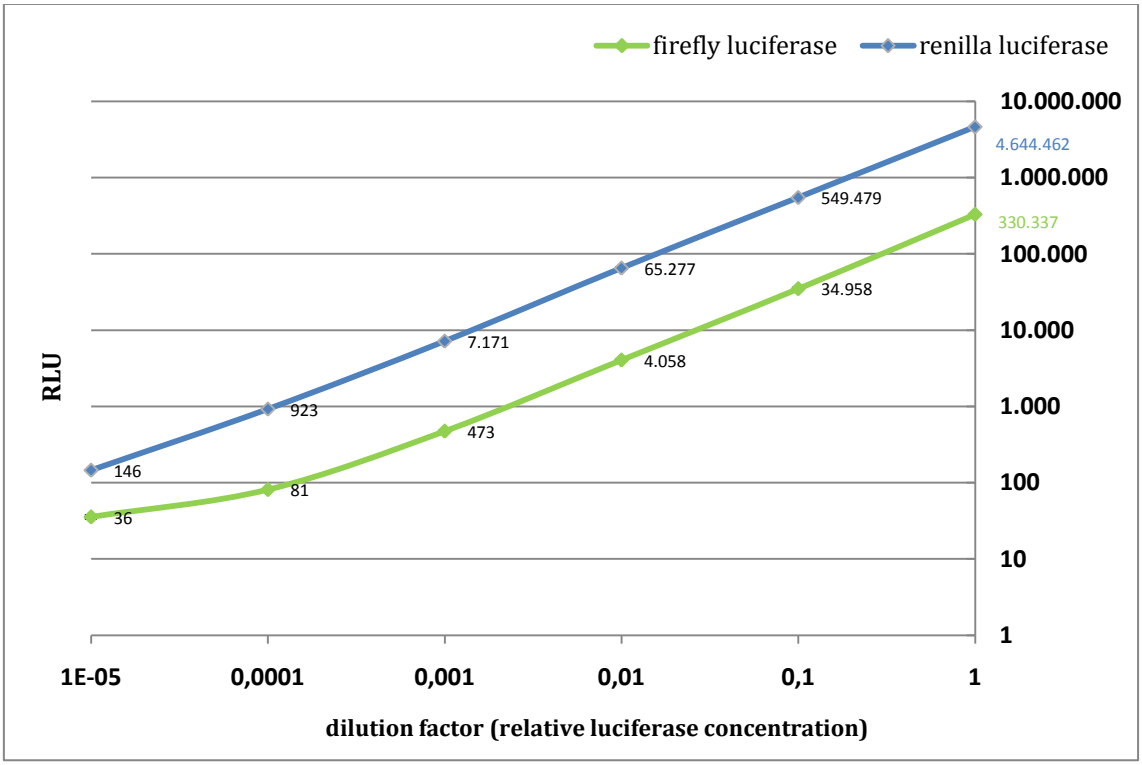


Fig. A3 **Linearity of RLU signal – lower RLU range.**

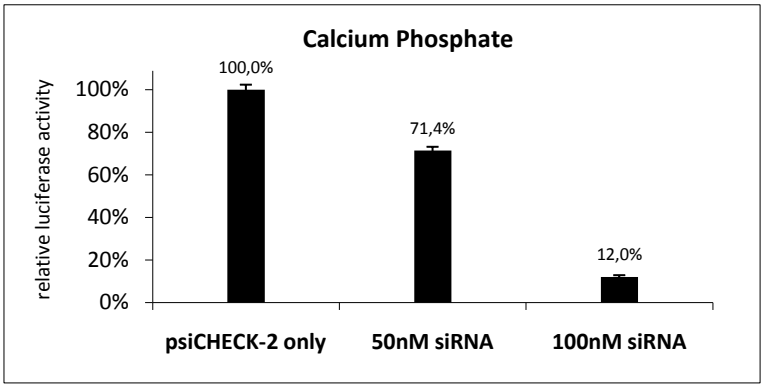


Fig. A4 siRNA dose dependency – CaP. Compared to DFD transfections (**Fig. R43**) calcium phosphate seems to either channel in more siRNA into the cytoplasm (where it is to mention that CaP transfections are conducted under RNA-typical conditions of ~50 to 60% confluence) or it shows its known strong protection of nucleic acids from nuclease attacks during lysosome passages. The error bars represent technical replicates. The drawback and exclusion criterion of CaP transfections is the unreliable reproducibility in biological replicates (though, caused by handling). *Setup:* 24-well format; 15.000 HEK293 cells seeded; transfection after 20 hrs according to Sigma's protocol; no glycerol shock; no serum during first 4 hrs of transfection; medium change after 12 hrs, *Medium I* applied; *psiCHECK-2* only: 0.5 µg *psiCHECK-2* + 0.5 µg *pBluescript* + 1.8 µl 2.5 M $CaCl_2$ in 18 µl sterile water; 50 nM: 0.5 µg *psiCHECK-2* + 0.67 µg *pBluescript* + 0.33 µg siRNA + 1.8 µl 2.5 M $CaCl_2$ in 18 µl sterile water; 100 nM: 0.5 µg *psiCHECK-2* + 0.34 µg *pBluescript* + 0.66 µg siRNA + 1.8 µl 2.5 M $CaCl_2$ in 18 µl sterile water; siRNA: *siGENOME* Non-Targeting siRNA Pool #2

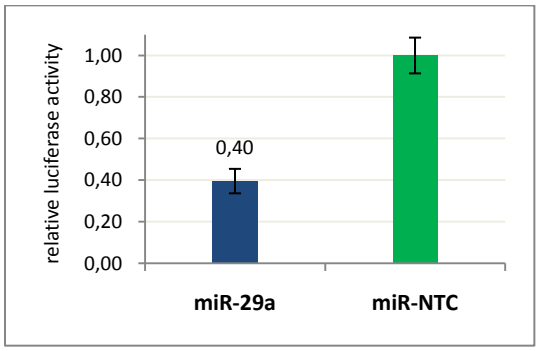


Fig. A5 Validation of miR-29a and TTP - CaP. 3.000 HEK293 cells per well of a 96-well plate were seeded 20 hrs before transfection, 90 ng of FL pGL3 construct (bearing either the two binding sites containing wildtype *WT* or double mutated *DM* form of the TTP 3' UTR) and 10 ng of RL pGL3 reference vector were co-transfected with 100 nM miR-29a or non-targeting control (NTC) miRNA. Cells were harvested 48 hrs post transfection. The data shown represent means±SD from two *biological* (independent) replicates. *Setup:* transfection according to Sigma's protocol; no glycerol shock; no serum during first 4 hrs of transfection; medium change after 12 hrs, *Medium I* applied; 0.28 µl 2.5 M $CaCl_2$ and 3.1 µl 2xHeBS pH7.05 per well

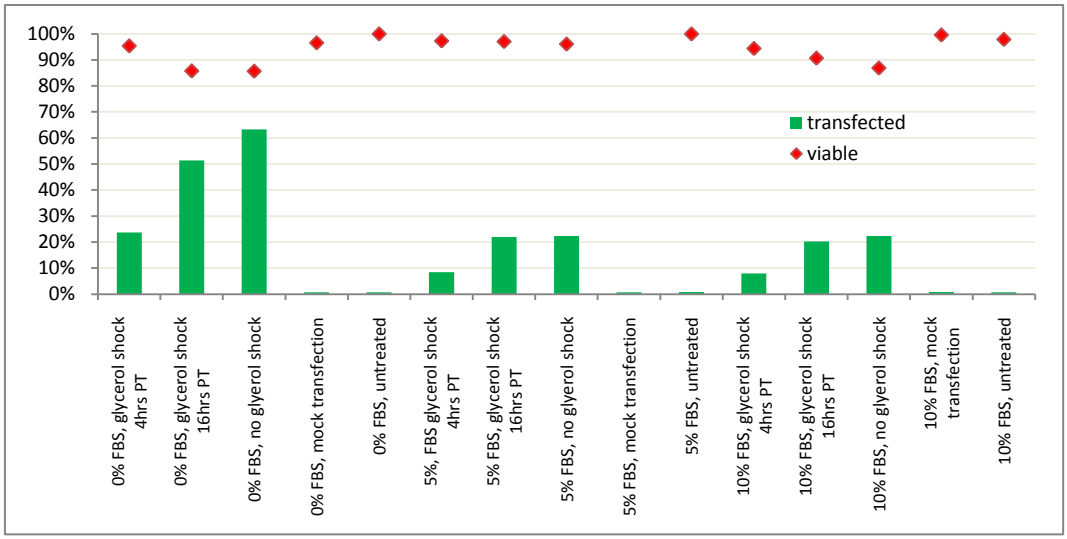


Fig. A6 Test for influences of FBS and glycerol shocking on CaP transfections. This test was the basis for the decision to do CaP transfections on HEK293 without serum and to leave complexes for 16 hrs on cells. *Setup*: 24-well format; 15.000 HEK293 cells seeded 20 hrs prior to transfection; transfection according to description in **Methods**; glycerol shocks 4 hrs PT: medium change to *Medium I* after 4 hrs in all other cases after 16 hrs;

Tab. A1 Primer sequences.

primer sequences	
PPARγ 3' UTR amplification forward	5'-CAT CAG CTC GAG CAG AGA GTC CTG AGC CAC T-3'
PPARγ 3' UTR amplification reverse	5'-CGG ATC GCG GCC GCA CTA TCA GCA ATT TCA TAA TAT GGT-3'
PPARγ 3' UTR site-directed mutagenesis primer forward	5'-ATT CTG AGG GAA AAT CTG ACA CCT AAG AAA TTT ACA CAC AAA AAG CAT TTT AAA AAG AAA AGG TTT TAG AAT AT-3'
PPARγ 3' UTR site-directed mutagenesis primer reverse	5'-ATA TTC TAA AAC CTT TTC TTT TTA AAA TGC TTT TTG TGT GTA AAT TTC TTA GGT GTC AGA TTT TCC CTC AGA AT-3'
psiCHECK-2 sequencing primer forward	5'-TAA GAA GTT CCC TAA CAC CG-3'
psiCHECK-2 sequencing primer reverse	5'-CGA GGT CCG AAG ACT CAT TTA G-3'

Tab. A2 siRNA sequences. (UU) = overhangs; sequences are for the sense strand

siRNA sequences (Dharmacon)	
siGENOME Non-targeting siRNA #2	5'-UAAGGCUAUGAAGAGAUAC(UU)-3'
siGENOME Non-targeting siRNA #4	5'-AUGAACGUGAAUUGCUCAA(UU)-3'

Tab. A3 PPARγ 3' UTR sequences and relevant miR sequences.

PPARγ and miR sequences	
PPARγ 3' UTR WT	CAGAGAGTCCTGAGCCACTGCCAACATTTCCCTTCTTCCAGTTGCACTATTCTGAGGGAAAATCTGACACCTAAGAAATTT **ACTGTGAA**AAAGCATTTTAAAAAGAAAAGGTTTTAGAATATGATCTATTTTATGCATATTGTTTATAAAGACACATTTACA ATTTACTTTTAATATTAAAAATTACCATATTATGAAATTGCTGATAGTA
PPARγ 3' UTR MUT	CAGAGAGTCCTGAGCCACTGCCAACATTTCCCTTCTTCCAGTTGCACTATTCTGAGGGAAAATCTGACACCTAAGAAATTT **AC<u>ACAC</u>AA**AAAGCATTTTAAAAAGAAAAGGTTTTAGAATATGATCTATTTTATGCATATTGTTTATAAAGACACATTTACA ATTTACTTTTAATATTAAAAATTACCATATTATGAAATTGCTGATAGTA
hsa-miR-27b MIMAT0000419	UUCACAGUGGCUAAGUUCUGC
Dharmacon microRNA Mimic Negative Control #1 (NTC)	UCACAACCUCCUAGAAAGAGUAGA
miR-NTC and miR-27b hybridization was checked against PPAR WT and MUT with rna22 (online tool); only miR-27b showed binding probability above threshold with settings: number of allowed UN-paired bases 0 in seed/nucleus of nucleotides 7; minimum number of paired-up bases in heteroduplex: 10; maximum folding energy for heteroduplex (kcal/mol): -15 **bold letters**: miR-27b binding site **underlined:** mutated bases	

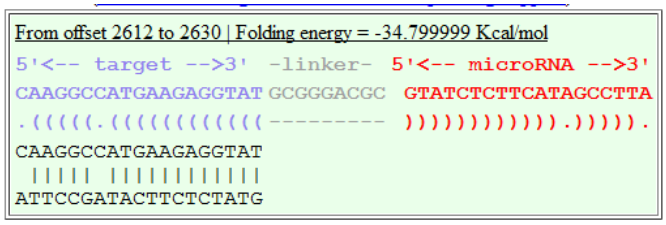


Fig. A7 Prediction for siRNA binding in the firefly luciferase. Sequence of siRNA#2 (see **Tab. A2**) was tested for binding in the firefly luciferase gene in the *psiCHECK-2* vector (AY535007) using *rna22* (online tool). siRNA#4 binding was not confirmed by *rna22*.

.

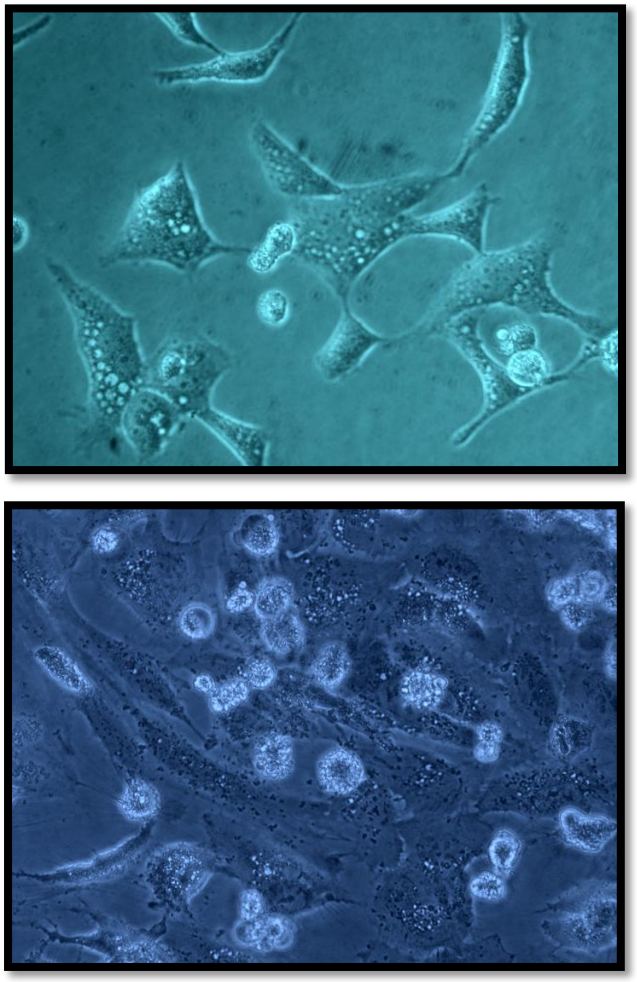


Fig. A8 Cellular stress caused by higher amounts of METAFECTENE Pro. Both, HEK293 (above) and HeLa (below) cells exhibit visually obvious cellular stress when facing high concentrations of MP. These pictures were taken from 1 µl MP transfections (96-well format).

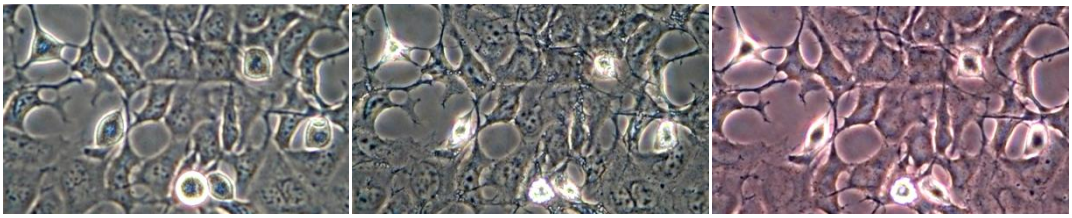


Fig. A9 Effect of a glycerol shock. (left) HEK293 cells before glycerol was added – watch the detached cells, (middle) osmotic shock: 1 min after addition of glycerol, (right) osmotic rescue: 1 min after PBS wash and addition of *Medium I*.

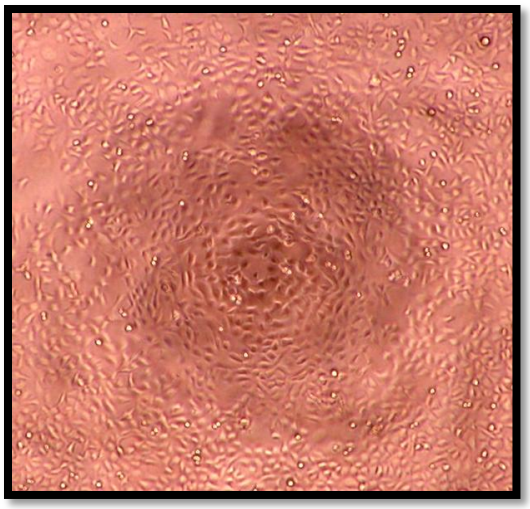


Fig. A10 HEK293 cell density at transfection. 20.000 cells seeded, 18 hrs post seed

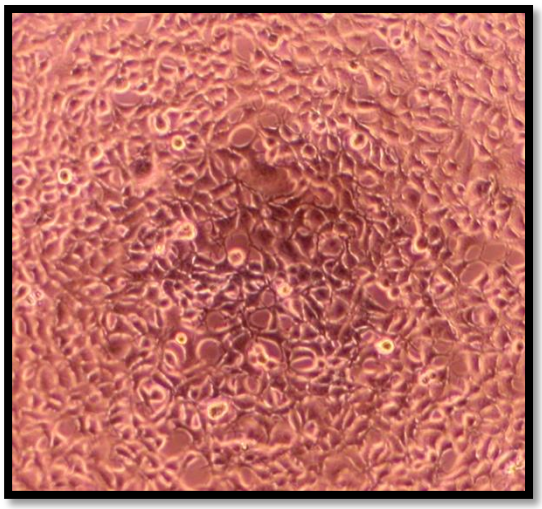
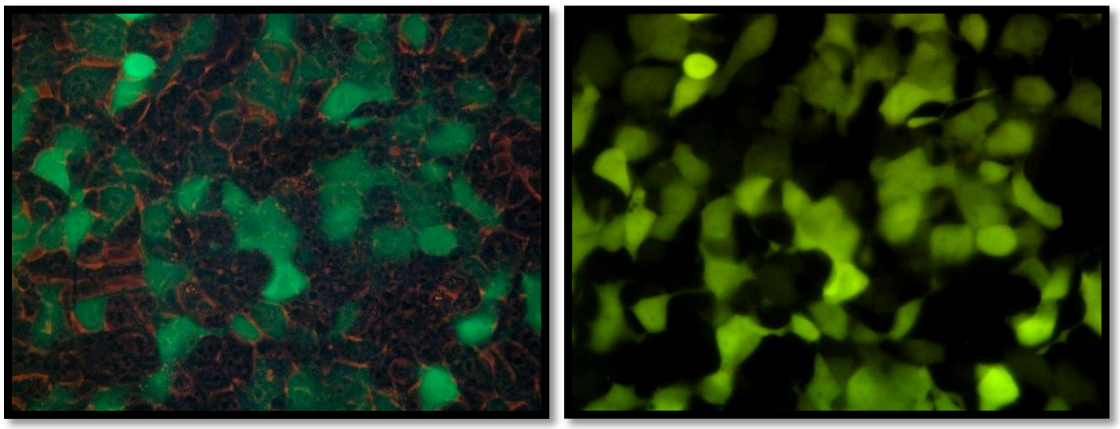


Fig. A11 HeLa cell density at transfection. 20.000 cells seeded, 18 hrs post seed



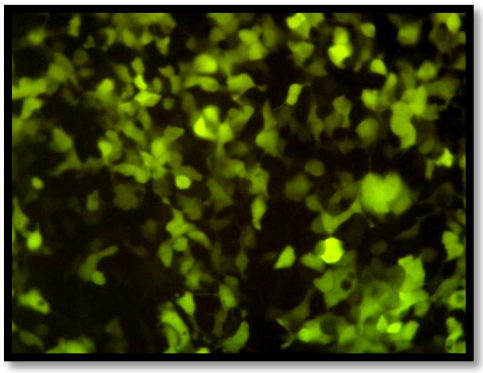


Fig. A12 Examples of eGFP transfected cells prior to FACS analysis. Shown is a sample with 75% eGFP expressing cells.

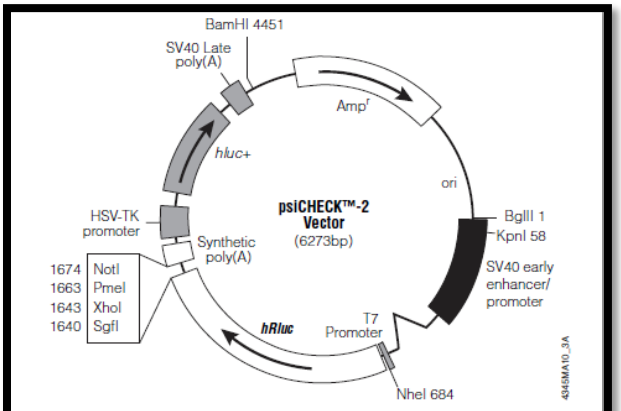


Figure 6. psiCHECK™-2 Vector map. -^- denotes the intron.

psiCHECK™-2 Vector sequence reference points:

SV40 early enhancer/promoter	7–425
Chimeric intron	489–621
T7 RNA polymerase promoter	666–684
Synthetic *Renilla* luciferase gene (*hRluc*)	694–1629
Multiple cloning region	1636–1680
Synthetic poly(A)	1688–1736
HSV-TK promoter	1744–2496
Synthetic firefly luciferase gene (*hluc+*)	2532–4184
SV40 late poly(A)	4219–4440
β-lactamase (Amp^r) coding region	4587–5447

Fig. A13 *psiCHECK-2* vector map