On Quantitative mRNA Transfection

Carolin Anke Leonhardt

München 2014

On Quantitative mRNA Transfection

Carolin Anke Leonhardt

Dissertation
durchgeführt an der Fakultät für Physik
der Ludwig–Maximilians–Universität
München

vorgelegt von
Carolin Anke Leonhardt
aus München

München, den 05.06.2014

Erstgutachter: Prof. Dr. Joachim O. Rädler
Zweitgutachter: Prof. Dr. Hermann E. Gaub
Tag der mündlichen Prüfung: 25.07.2014

Contents

List of Figures

Zusammenfassung

Das künstliche Einschleusen genetischer Information in Form von exogener Plasmid Desoxyribonukleinsäure (pDNA) oder messenger Ribonukleinsäure (mRNA) in Zellen bezeichnet man als *Transfektion*. Infolge der Transfektion werden in der Zelle neue Proteine nach dem Bauplan des auf der pDNA/mRNA kodierten genetischen Codes exprimiert. Es zeigt sich, dass dabei starke Unterschiede im zeitlichen Verlauf der Proteinexpression und in der Anzahl exprimierter Proteine pro Zelle auftreten. Diese Phänomene waren bislang ungeklärt.

Im Rahmen der vorliegenden Arbeit wurden Messungen zum zeitlichen Verlauf der Expression solch künstlicher Gene durchgeführt und eine stochastische Beschreibung der mRNA Transfektion entwickelt. In den Experimenten wurden mit Hilfe automatisierter Langzeit-Fluoreszenzmikroskopie viele tausend Zellen pro Messung abgerastert und dabei die Zeitverläufe der Fluoreszenzintensität jeder individuellen Zelle aufgenommen. Die kontinuierliche Beobachtung individueller Zeitverläufe ist entscheidend, um die zugrunde liegende Variabilität von Zelle zu Zelle zu erfassen, da diese Information in herkömmlichen Ensemble-Messungen ausgemittelt wird.

Im ersten Teil wurden eukaryontische Zellen mittels kationischer Lipide mit *in vitro-transkribierter (IVT)* mRNA oder pDNA transfiziert, welche für das Reporterprotein EGFP (engl. *enhanced green fluorescent protein*) kodieren. Es zeigte sich, dass die Zeitverläufe der mRNA-vermittelten EGFP Expression durch die analytische Lösung eines einfachen biochemischen Ratenmodells für Translation und Degradation beschrieben werden kann. Auf diese Weise wurde die Verteilungsfunktion der mRNA Degradationsraten ermittelt. Die Anzahl fluoreszenter Proteine pro Zelle sowie der Prozentsatz erfolgreich transfizierter Zellen wurden zusätzlich mit Hilfe von Durchflusszytometrie bestimmt. Im Vergleich zu den langen *onset*-Zeiten von bis zu 20 Stunden, die für pDNA Transfektion gemessen wurden, findet man für die mRNA Transfektion kurze *onset*-Zeiten zwischen 0.5 und 5 Stunden. Dies ist in Übereinstimmung mit der Tatsache, dass bei mRNA Transfektion der raten-limitierende Schritt der pDNA Transfektion, nämlich der Eintritt in den Zellkern, entfällt. Auf Basis einer zusätzlich gemessenen Dosis-Wirkungs-Kurve wurde ein Zwei-Schritt mRNA Transfektionsmodell aufgestellt, welches Aussagen über die Anzahl effektiv transfizierter Komplexe aus mRNA und Helferlipid (Lipoplexe) erlaubte. Auf Grundlage dieses Modells konnte die Verteilung der Expressionsraten sehr gut beschrieben und erstmals eine Erklärung für die Varianz der Expressionsniveaus gegeben werden.

Ausgehend von der Möglichkeit, auf oben beschriebene Weise mRNA Lebenszeiten zu messen, wurden in Zusammenarbeit mit der Arbeitsgruppe von Professor C. Plank (Technische Univer-

sität München) in einem weiteren Projekt die Verteilungen der Lebenszeiten von sechs chemisch modifizierten mRNA Konstrukten gemessen, die verschiedene Kombinationen von stabilisierend wirkenden untranslatierten Bereichen (engl. *untranslated regions (UTRs)*) eines humanen Gens aufwiesen. Um eine ausreichende Statistik im Hochdurchsatz generieren zu können, wurden die Zellen hierzu auf mikrostrukturierten Substraten kultiviert. Dies ermöglichte eine nahezu vollautomatisierte Bildverarbeitung der Einzelzellspuren. In zwei verschiedenen Zelltypen wurden maximal lebenszeitsteigernde Wirkungen für solche Konstrukte gefunden, die entweder nur an ihrem 3'-Ende, oder aber sowohl am 5'- als auch am 3'-Ende UTRs enthielten. Durch Integration der analytischen Lösung der mRNA-vermittelten Proteinexpression für $t \to \infty$ konnte eine wichtige pharmakokinetische Kenngröße, die Integralfläche der Dosis-Zeit-Kurve (engl. *area under the curve, AUC*), extrapoliert werden.
Schließlich wurde auf Grundlage experimenteller Daten ein Multi-Level Modell für mRNA Transfektion aufgestellt, mit dem die Kinetik des Gentransfers und damit die experimentell bestimmte *onset*-Zeitverteilung reproduziert werden kann. In einem iterativen Prozess wurden durch stochastische Simulation generierte Daten analog zu reellen Daten gefittet. Die daraus resultierenden Verteilungen wurden mit Hilfe eines Standardverfahrens für globale Optimierung an die von mir experimentell bestimmten Werte angepasst. Das Modell konnte auch experimentelle Dosis-Wirkungskurven reproduzieren sowie den Einfluss bestimmter Raten auf die Transfektionseffizien vorherbestimmen.

Abstract

In transfection experiments, exogenous plasmid deoxyribonucleic acid (pDNA) or messenger ribonucleic acid (mRNA) are artificially introduced into cells. Subsequently, new proteins are expressed according to the genetic information that is encoded on the nucleic acid. To date, a quantitative description of the temporal cell-to-cell variability of the resulting protein expression and the differences in numbers of expressed proteins is missing.

This thesis is focussed on measuring the time courses of gene expression. In addition, a new model for mRNA transfection that relates experimental observations to a stochastic description of mRNA transfection was developed. Only continuous, single-cell observation is able to elucidate cell-to-cell variabilities, which are averaged out in ensemble measurements. Therefore, most experiments of this thesis consisted of automated, long-term fluorescence time-lapse microscopy. This way, thousands of cells were scanned in each measurement and the fluorescence intensity time courses of individual cells were monitored.

The first part of this thesis reports on a quantitative study about transfection of eucaryotic cells with *in vitro-transcribed (IVT)* mRNA or pDNA. Both of them encode for the enhanced green fluorescent protein (EGFP) as a reporter and were introduced to the cells by cationic helper lipids. We were able to fit the time courses of mRNA-mediated protein expression using the solution of a simple biochemical rate model for translation and degradation. This way, the distribution function of mRNA life times could be determined. Furthermore, the number of fluorescent proteins per cell and the percentage of successfully transfected cells were measured by flow cytometry. Consistent with the fact that mRNA transfection lacks the rate-limiting step of nuclear entry, the onset times ranged from 0.5 h to 5 h only. In contrast, longer onset times from 2 h to 20 h were observed for pDNA transfection. Based on an experimental dose-response curve, a simple two-step model for mRNA tranfection was developed. This enabled us to estimate the number of effectively delivered lipoplexes (i.e. the complexes of condensed mRNA molecules and cationic helper lipids) and to describe the distribution of the expression rates. Moreover, the variability of the protein expression levels can be explained by our model.

The method described above to measure mRNA life times was also used in a second project, which was carried out in close collaboration with the group of Professor C. Plank (Technische Universität München). Here, the stability of six different, chemically modified mRNA constructs was assessed. The constructs differed in the number and combination of inserted *untranslated regions (UTRs)* of a human gene which are known to cause stabilising effects. However, to be able to generate

decent statistics at high throughput, cells were grown on micropatterned substrates to allow for automated image analysis of single-cell time courses. The best life time prolongation was detected for constructs that had UTR insertions either at their 3' end or at their 5' and 3' ends in two different cell types. Interestingly, by integration of the analytical solution for mRNA-induced protein expression for $t \to \infty$, we were able to extrapolate the *area under the curve (AUC)*, which is a key pharmacokinetic parameter.

Finally, a multi-level kinetic model for mRNA transfection and the kinetics of gene transfer was developed on the basis of experimental data. This model correctly reproduced experimental dose-response curves as well as onset time distributions. The data obtained by stochastic simulation were analysed in exactly the same way as the corresponding experimental data and the resulting distributions were iteratively adjusted to the parameter distributions of the experimental data sets using a standard technique for global parameter optimisation. In the spirit of predictive modeling, the effect of specific rates on the resulting transfection efficiencies was analysed.

1 Introduction

Various revolutionary discoveries and technical developments of the past decades have paved the way for measuring and monitoring biochemical reactions and processes in the natural environment of live cells. In fact, living cells can be considered to be "the test tubes of the 21st century", as claimed by Jonathan Widon in his famous remark[1]. Achievements such as the discovery and subsequent engineering of a multitude of fluorescent proteins used as reporters [1–6], super-resolution microscopy [7–10], and fluorescent labeling techniques as well as single-particle tracking (SPT) approaches [11–14] allow the study of individual molecules with unprecedented temporal and spatial resolution inside living organisms [15–18].

The field of gene delivery

The field of molecular medicine and gene therapy has evolved rapidly and the vision of personalised and tailored drugs that treat diseases such as cancer at targeted sites of action is becoming increasingly realistic [19]. For example, more than 900 clinical trials involving viral gene transfer were conducted between 1989 and 2004 [20]. Following the spirit of the "magic bullet" as proposed by the Nobel Prize laureate and inventor of chemotherapy, Paul Ehrlich, more than 100 years ago [21], a multitude of viral and non-viral delivery systems for drug delivery have been proposed in the last years [22–24]. Due to the high safety risks that are associated with viral gene delivery, a number of improved, highly efficient and non-toxic non-viral alternatives have been developed. These include chemical as well as physical methods, such as cationic lipids, cationic polymers, or magnetic carriers [25–38]. Most recently, DNA-based nanostructures acting as carrier systems have also been introduced and shown promising results for the innovative field of nanomedicine [39, 40]. Besides the transfer of actual drugs into cells, for another class of applications cargos such as pDNA or RNA molecules that induce the production of therapeutic proteins inside the cell are of high importance. pDNA transfection generally suffers from poor transfection efficiencies that stem from the numerous intracellular barriers. The discovery of small interfering RNAs (siRNA) introduced an entirely new class of molecules to be delivered in gene-therapy applications, with the possibility of suppressing the expression of malignant proteins by silencing specific mRNA sequences inside target cells [42, 43].

[1] Jonathan Widom, 1955-2011

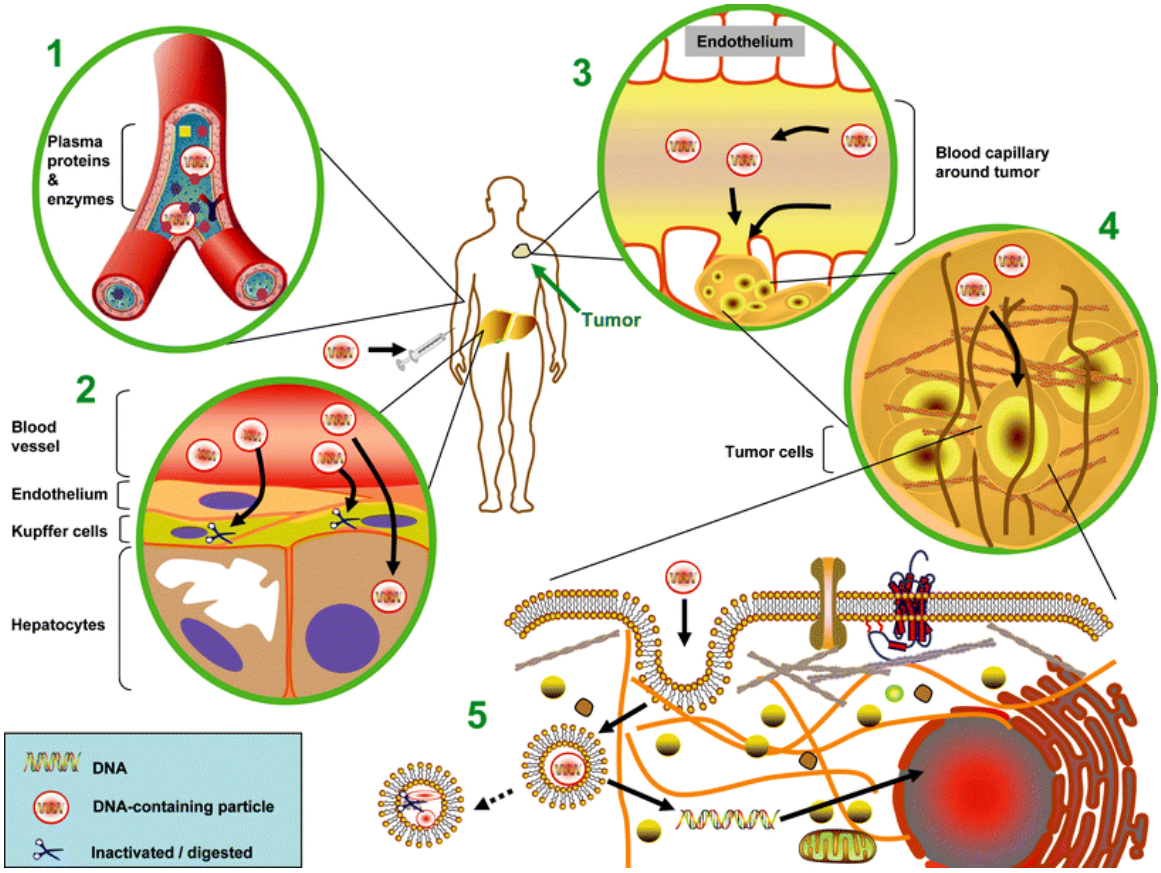

Figure 1.1: *Intravenous, liposome-aided gene delivery and its barriers. (1) After intravenous injection, serum proteins may bind to cationic lipoplexes and lead to particle aggregation. (2) The aggregates can be taken up by macrophages and thus be eliminated from circulation. (3) Part of the circulating lipoplexes may extravasate into tumor tissue. (4) Lipoplexes may manage to transverse the dense extracellular matrix and reach the cell surface of the target tumor cells. (5) After cellular uptake, the lipoplex cargo has to successfully complete the pathway of intracellular gene delivery in order to come into effect. Figure taken from [41] and reprinted with permission from Springer, copyright 2007.*

The role of mRNA for gene delivery

In gene delivery, mRNA has become increasingly popular as a carrier of genetic information to induce transient gene expression without causing genetic modifications. The first non-viral transfer of mRNA into cells was carried out in 1989 by Malone, Felgner, and Verma, and first clinical studies of mRNA therapeutics were conducted in the late 1990s [44, 45]. Today, mRNA is one of the upcoming approaches for gene therapy applications, not only because it circumvents the risk of mutagenesis [45–50].

The introduced mRNA can encode for any protein and its transfection is dose-titratable. The high efficiency of mRNA transfection offers the possibility to simultaneously co-transfect different therapeutic mRNAs at tunable doses. Unlike pDNA, mRNA does not carry the risk of chromosomal integration of transgenes by the target cell. Therefore, it is better-suited for clinical applications. Further advantages of mRNA-induced protein expression as compared to pDNA are the short onset

times of protein expression and its efficacy in non-dividing cells [51–54]. However, transfection of mRNA into mammalian cells can cause severe cytotoxic effects due to innate immune response by Toll-like receptors and RNA sensors [55–57]. Also, first artificial mRNA constructs suffered from low stability. These problems have been overcome by the development of long-lived, non-immunogenic mRNAs that further underline the importance of mRNA for gene-therapy applications [49, 58, 59]. For example, the localised use of mRNA by pulmonary delivery has recently been demonstrated in mice [49] and promising results of mRNA transfection for dendritic cell-based immunotherapy have been obtained [60, 61].

The discovery that differentiated cell types can be reprogrammed to a state of pluripotency was a monumental moment in stem cell biology and the associated field of regenerative medicine [62, 63]. Most recently, modified mRNA has been used by several groups for reprogramming of human fibroblasts [64–66], which is very promising for regenerative therapies. Here, mRNA transfection acts as a highly-controllable tool. It has also been used for multiple transfection cycles of mouse and human neural stem cells to achieve sustained protein expression while avoiding the disadvantages of pDNA transfection [53].

How to measure mRNA abundance in cells

Today, many efforts are on the way to investigate the role of mRNA as a key player in regulated gene expression, where it facilitates coordinated gene expression, and the mRNA-decay machinery [67]. Because mRNA turnover plays such an essential role in the coordinated network of regulated gene expression, methods to measure mRNA abundance as well as the corresponding rates without perturbing the cellular system are urgently needed.

Classically, experiments to measure mRNA levels use Western blot or polymerase chain reaction (PCR)-based methods such as reverse-transcription PCR (RT-PCR) [68], real-time PCR (qPCR) [69], or real-time RT-PCR (qRT-PCR) [58, 70, 71]. Also, the rates of mRNA decay can be measured at the whole transcriptome level using transcription blocking agents and microarrays for time resolved monitoring of mRNA decay [72]. To quantify the amount of newly transcribed mRNA, metabolic labeling is often used, where the nucleoside analogue 4 - thiouridie (4 sU) is taken up by cells and incorporated into newly transcribed RNA, which can then be isolated and quantified [73–75]. In 1998, Femino et al. introduced the technique of fluorescence in situ hybridisation (FISH), which can count individual mRNAs within single cells and has be used in combination with SPT to follow the life cycle of mRNA in live cells [12, 13]. However, FISH relies on cell fixation protocols and requires many probes taken per molecule for a detectable signal. Also, its temporal resolution is not sufficient to accurately study dynamic cellular processes. A new tool for measuring mRNA synthesis and decay rates in yeast on a global scale and in a dynamic manner without perturbing the system has been presented recently [76]. The results obtained by this ”dynamic transcriptome analysis” (DTA) suggested that in yeast, mRNA decay and synthesis are functionally independent during normal growth and that both processes contribute to setting cellular mRNA levels [76]. Most recently, microwell-based RNA cytometry has been employed to reveal cellular heterogeneity of stem cell populations [77]. These methods stand in contrast to the highly time-resolved, indirect

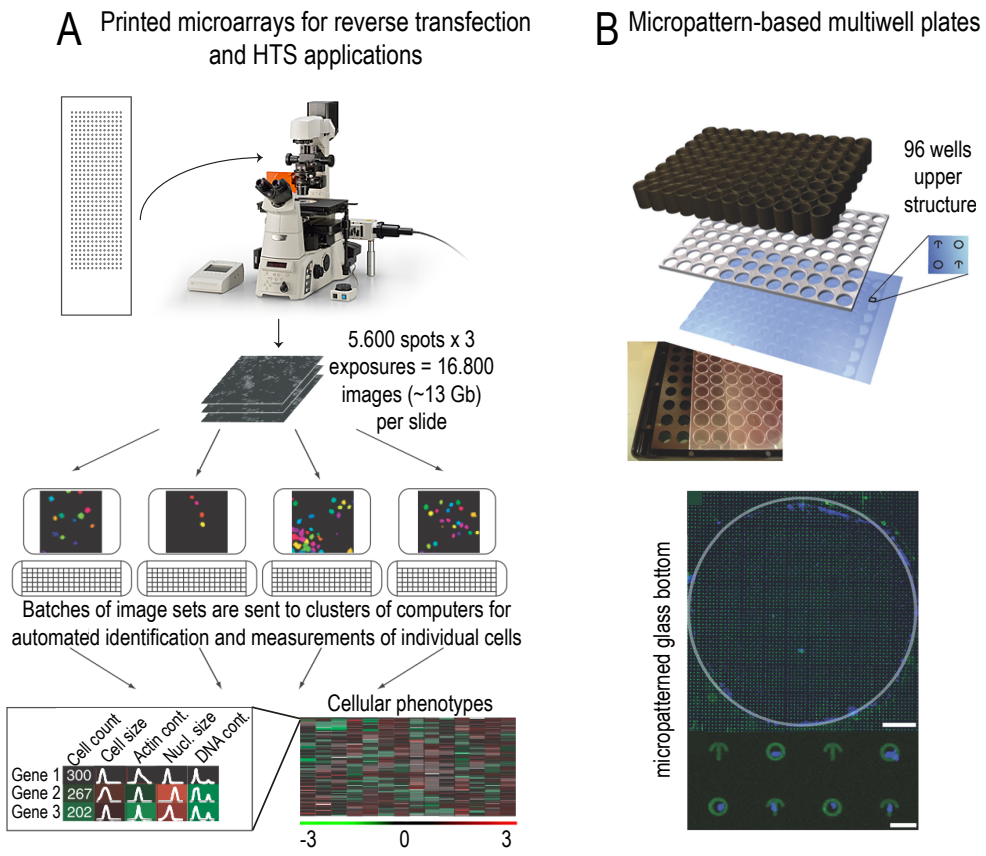

Figure 1.2: *Controlled microenvironments for high-throughput applications. (A) Printed microarrays for reverse siRNA transfection in combination with automated microscopy can be used for loss-of-function screens of entire RNA libraries [78]. After incubation with cells, such microarrays can be fixed, stained and imaged for analysis. Alternatively, they can be imaged live. Parameters such as reporter expression, cell viability, cell number, or cell morphology can be analysed. Adapted by permission from Macmillan Publishers Ltd: Nature Genetics [78], copyright 2005. (B) Piel et al. developed a simple and robust process for patterning glass substrates using deep UV light [79]. This technique can be used to micropattern multiwell plates with proteins to promote adhesion of single cells. The high number of confined cells makes this technique a powerful tool for high-throughput screening applications and standardised cell assays. Scale bars* $40\,\mu m$*. Adapted from [79] with permission of The Royal Society of Chemistry.*

method of determining mRNA levels by fitting single-cell time courses of EGFP expression levels, which will be presented in this thesis.

Cell arrays and the importance of single-cell observation

Time-lapse fluorescence microscopy, which makes use of fluorescent reporter proteins, is a standard procedure in current research in cell biology, cell biophysics and medicine [80–89]. In combination with micropatterned substrates (see Figures 1.2 and 1.5) which array the cells in tunable arrangements, time-lapse fluorescence microscopy can be used to measure single-cell time courses [84, 90]. To promote ordered cell growth, various micropatterning techniques have been developed.

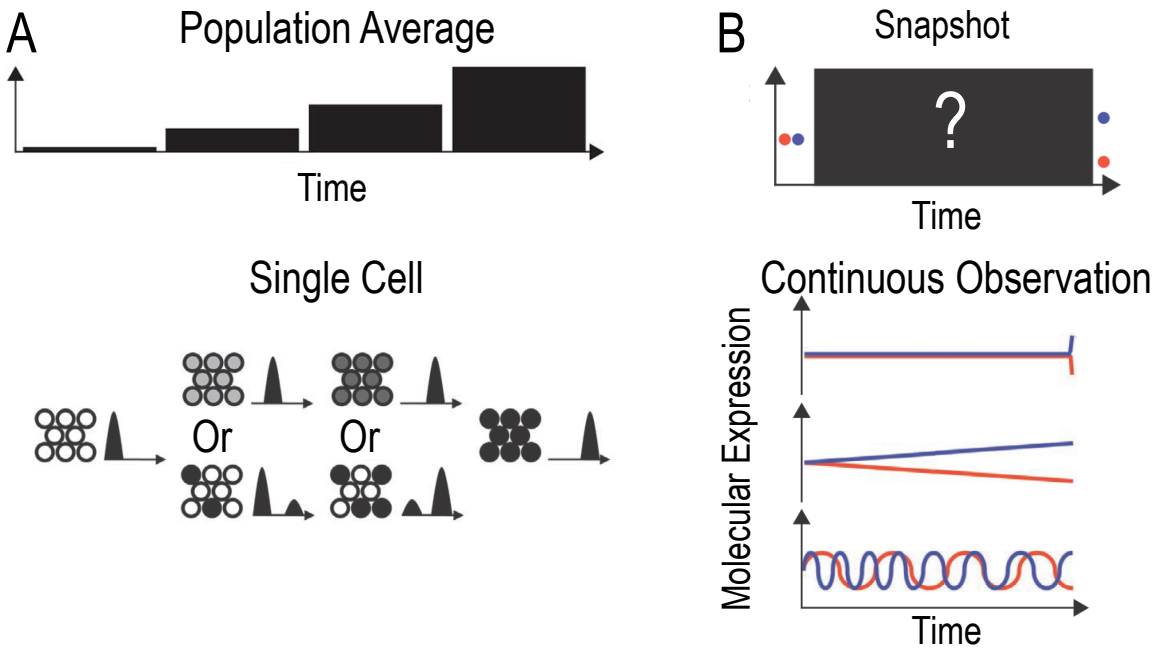

Figure 1.3: *The importance of continuous observation at the single-cell level: (A) Ensemble measurements obscure single-cell variability inside clonal populations. (B) In contrast to snapshot measurements, continuos observation allows monitoring the temporal evolution of expression patterns inside cells. Adapted by permission from Macmillan Publishers Ltd: Nature Methods [89], copyright 2011.*

Arraying cells facilitates automated image analysis, which eliminates the bias of visual inspection and significantly speeds up image analysis. Such experimental setups close the gap between traditional microscopes and plate readers and are a promising approach for standardised assays and high-throughput cell studies such as loss-of function screens using RNA interference (RNAi) (see Figure 1.2) [78, 79]. While many of the current high-throughput applications yield ensemble values, single-cell measurement techniques are needed because population-based averages of cell-biological parameters are inherently unable to uncover the underlying cell-to-cell variability [69, 86, 91, 92]. To profoundly understand the stochasticity of gene expression and the resulting consequences that have been reviewed in [93], single-cell observation is indispensable. In fact, only continuous observation at the single-cell level allows simultaneous detection of temporal parameter variations and their distribution amongst a population of genetically identical cells, as illustrated in Figure 1.3 [81, 84, 85, 89, 94].

Predictive modeling and its relevance for synthetic biology

With the increasing number of non-viral vectors for gene delivery that have been developed over the past years, the need for models that are based on quantitative experimental techniques has become evident. Combining quantitative experimental data with computational modeling facilitates the process of getting a precise understanding of the kinetics of the administered gene. Models that represent the cellular and intracellular delivery of exogenous genes as comprehensively and correctly as possible are essential tools for the analysis, characterisation and optimisation of synthetic

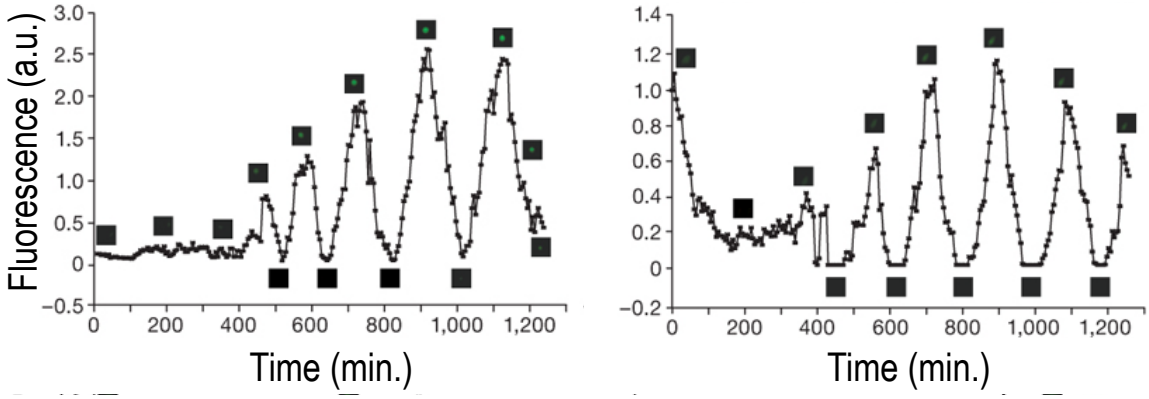

Figure 1.4: *A combined experimental and modeling approach was used by Tigges et al. to create a tunable, self-sustained synthetic oscillator in eucaryotic cells [95]. Two exemplary single-cell time courses are shown here to illustrate the oscillating GFP levels that were induced by the synthetic gene network and monitored using time-lapse fluorescence microscopy. Adapted by permission from Macmillan Publishers Ltd: Nature [95], copyright 2009.*

vectors and might provide a theoretical foundation for developing products with tailored therapeutic profiles. However, each vector has its unique uptake and intracellular trafficking kinetics. Despite optimistic theoretical and experimental predictions for a gene delivery system, additional *in vivo*-barriers (as shown illustrated in Figure 1.1) might still come into play when a vector is finally tested in clinical trials.

A quantitative understanding of the gene-delivery process is also needed for the further development of synthetic devices that rely on cell transfection with exogenous genes as, for example, proposed in [95, 96]. Figure 1.4 shows two exemplary single-cell time courses of the synthetic mammalian oscillator that was introduced by Tigges et al. [95]. Here, three interplaying pDNA constructs were transfected into Chinese hamster ovary cells and time-lapse fluorescence microscopy was combined with mathematical modeling. Interestingly, differences in the timing of pDNA uptake (*i.e.* the gene transfer process) resulted in cell-to-cell variability of GFP expression. Understanding such synthetic gene circuits will hopefully advance the understanding of complex intracellular control networks and signaling pathways that are also involved in severe diseases.

The pharmacokinetics of gene therapies were first described by a numerical multicompartment model of cellular events in 1994 by Ledley et al. [97]. Soon after, electron microscopy was used to elucidate the pathway of liposomes inside cells [98]. Varga and co-workers were the first to present a kinetic model that focused on the characterisation of gene delivery processes [99]. This model correctly predicted the dependence of transgene expression on the length of the cationic polymers that were used for transfection as well as a significant increase in gene expression only above a threshold unpackaging rate. However, this framework was based on mainly polyplex-based transfection of pDNA only. Because the study founded its basis in snapshot measurements, the authors pointed out the need for dynamic experimental data. Soon after, they reported on a modification of this mathematical mass-action model for the comparison of quantitative parameters that characterise intracellular gene delivery steps and a sensitivity analysis of different vectors [100]. This way, they

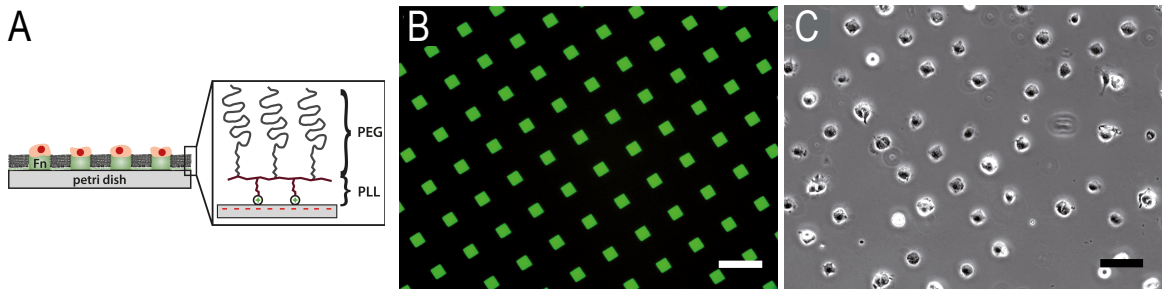

Figure 1.5: *The micropatterning technique described in [90] was used to create single-cell arrays for the measurements of chapter 4. (A) Cell-adhesive islands and cell-repellent areas promote ordered cell growth. (B) Fluorescently labeled fibronectin was used to visualise the regular pattern. (C) Cells self-assemble onto the adhesive islands after seeding. Scale bars* $100\,\mu m$*. Adapted from [90] with permission of The Royal Society of Chemistry.*

were able to elucidate which steps are faster for a superior vector as compared to an inferior vector. A pharmacokinetic model to analyse the rate-limiting process of non-viral gene delivery was developed by Harashima et al. [101, 102]. Similar to the work of Varga, their optimisation strategy for intracellular trafficking of exogenous genes also focussed on pDNA. In another study, the effect of the different intracellular composition on kinetic rates was investigated on the basis of a mathematical model [103]. However, kinetic models approximate all spatial and transport processes by kinetic equations. A computational framework for intracellular trafficking of polyplexes that takes the cell size and shape into account could qualitatively predict the spatiotemporal distribution of polyplexes inside cells as determined by SPT [104]. The need for such models that combine spatial and stochastic aspects to comprehensively represent the dynamics of biochemical systems has been pointed out recently [105]. For further advancement of predictive models, experimental techniques that allow quantitative and direct measurement of actual molecule numbers will be required [106]. In addition to the models discussed here, statistical frameworks for the uptake of nanoparticles and pDNA-lipoplexes have been developed recently [88, 107]. Still, a general problem of existing delivery models is the fact that they do not cover effects such as loss of pDNA or mRNA by dilution in dividing cells, which is a general problem in existing models.
Besides the above-mentioned delivery models, stochastic, quantitative and kinetic models for gene expression exist [75, 108, 109]. Also, different components of noise in gene regulatory networks could be identified on the basis of single-cell measurements and stochastic models [94, 110]. In 2008, a two-stage model for gene expression in budding yeast was proposed, where analytical expressions were used to avoid large numbers of simulation runs. This approach exploited the differences in timescales, thus simplifying the dynamics of the system [111].

This work

Quantitative experimental studies in coordination with mathematical modeling are able to provide detailed, quantitative understanding of non-viral gene transfer, as will be demonstrated in this thesis.

Remaining challenges such as measuring the distribution functions of mRNA and protein life times, as well as mathematically describing the stochastic processes that are involved in artificial gene delivery and expression will be tackled in this thesis. To this aim, a comparative, yet quantitative study on mRNA and pDNA transfection that was analysed at the single-cell level, as well as a two-step stochastic model for mRNA delivery are presented in chapter 3. Here, we determined single-cell distributions of characteristic kinetic rates such as the mRNA degradation rate or the onset times of protein expression and related the distribution of the protein expression levels to a mathematical model. The same approach was combined with microstructured single-cell arrays (see Figure 1.5) to measure the life times of differently stabilised mRNA constructs in high-throughput. This project is described in chapter 4. A multi-level kinetic model for mRNA transfection that has been developed on the basis of experimental data and which we used to make predictions is presented in chapter 5.

2 Basic concepts, experimental techniques and image analysis

2.1 pDNA and mRNA as carriers of genetic information

For most experiments of this thesis, fluorescent reporter proteins were used to monitor protein expression. To induce the expression of such foreign proteins in cells, molecules carrying the corresponding genetic code have to be introduced into the cells for subsequent transcription and/or translation into the protein. In general, such molecules can be either artificial pDNA or mRNA.

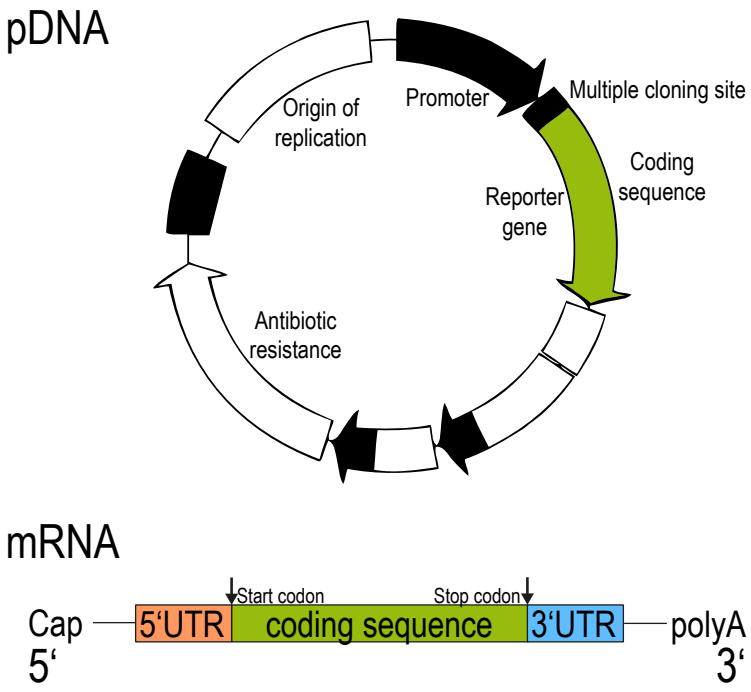

Figure 2.1: *Schematic representations of a pDNA vector and a typical mature mRNA molecule. Artificial mRNA vectors as shown above can be produced from pDNA using commercial in vitro transcription kits.*

pDNAs are small, circular, double-stranded DNA molecules that can replicate independently of chromosomal DNA. The essential parts of a pDNA vector are the origin of replication, the promoter region, the coding sequence, and antibiotic resistances. The DNA sequence consists of four

nucleobases that encode for the mRNA sequence. To convey genetic information from DNA to the ribosomes, mRNA is transcribed from DNA. During translation, ribosomes translate the amino acid sequence of the protein from the mRNA. Both artificial pDNA and mRNA are often used for molecular cloning and have been used for the transfection experiments of this thesis. To this end, mRNA was *in vitro* transcribed. Artificial mRNA vectors that are produced in this way typically consist of a 5' cap, a 5' untranslated region (UTR), a coding sequence, a 3' UTR and a polyA tail. UTRs and polyA tail are known to be important for mRNA stability [58, 112].

2.2 The general concept of transfection

Transfection is the process of introducing foreign nucleic acids into eukaryotic cells. It is a common approach to study protein expression or gene function and to transfer DNA or RNA into various types of cells, including stem cells [47, 63]. Since the first report of artificial gene transfer into procaryotic cells (termed *transformation* for bacteria) in 1973 [113], various methods to transfect also eukaryotic cells have been developed.

In general, one can choose between viral and non-viral methods. In viral gene delivery, which is also referred to as *transduction*, viruses are employed to transfer genes into target cells [114, 115]. However, viral vectors carry a high risk of causing host immune response and insertional mutagenesis.

As an alternative, there is a broad variety of non-viral gene delivery vectors that can act as artificial viruses. For example, nucleic acids can be transferred into cells by microinjection [25], by magnetofection [33, 36], or by electroporation [26]. Also, polymers [34, 37, 38], dendrimers [28, 35], cell-penetrating peptides [32], calcium-phosphate [29], or liposomes [27, 30, 31, 116, 117] can be used for transfection. Another approach for gene delivery is the use of stealth liposomes [118–121]. For all transfection experiments that are part of this thesis, the liposomal approach was chosen. More precisely, the commercial, liposome-based transfection reagent *Lipofectamine 2000* ® was used. Transfecting cells by means of cationic lipids in the form of liposomes, a method which is also referred to as *lipofection*, is a very common and reliable way of transfection, resulting in high transfection efficiencies.

Figure 2.2 shows a simplistic illustration of lipofection. Negatively charged nucleic acids (either RNA or pDNA) are incubated with cationic lipids and self-assemble into complexes, the so-called *lipoplexes*. Both structure and density of these lipoplexes have been studied before using small angle x-ray scattering (SAXS), x-ray diffraction, and atomic force microscopy (AFM) [122, 123]. After sedimentation, the lipoplexes attach to the negatively charged cell membrane. Because their overall surface charge is positive, they bind to the proteoglycans via electrostatic interaction [124, 125]. Alternatively to this top down approach, complexes can also be presented to the cell on the substrate itself before cell seeding. This transfection method is called *reverse transfection* [126, 127]. Regardless of the way complexes reach the cell surface, they are subsequently incorporated by the cells via endocytosis [128]. After uptake, RNA-lipoplexes can be unpacked and translated into protein directly after endosomal release into the cytosol. In contrast to that, the DNA load of

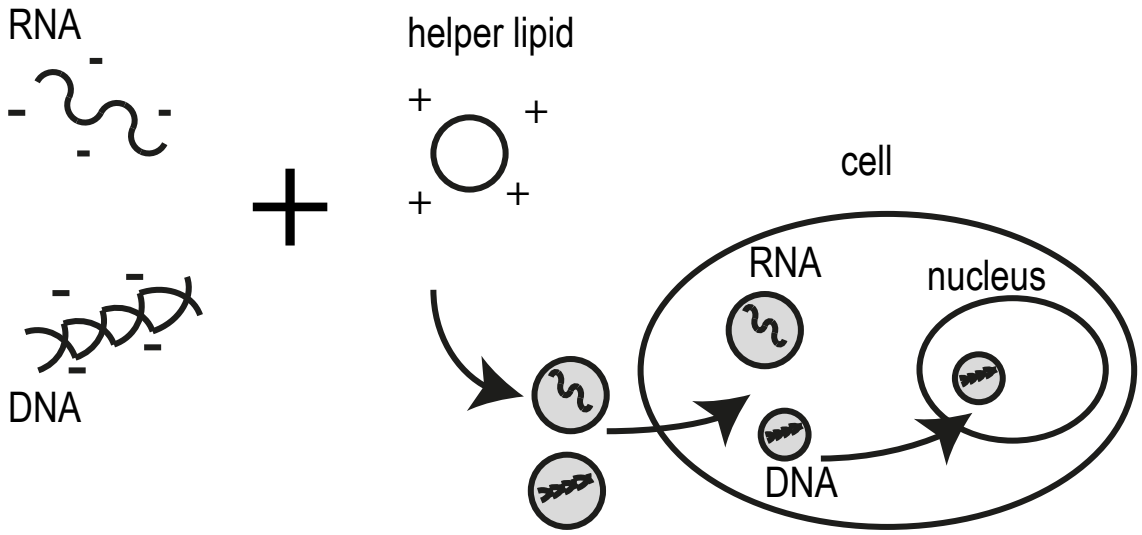

Figure 2.2: *Basic principle of lipofection. The anionic nucleic acid (pDNA/mRNA) is complexed with a cationic helper lipid and taken up by the cell via the endocytotic pathway. While RNA can be directly processed in the cytoplasm, DNA has to enter the nucleus in order to be transcribed into mRNA.*

pDNA-lipoplexes first has to enter the nucleus where it can be transcribed into mRNA. This event is more likely to occur when the nuclear membrane disappears during mitosis [129]. Still, the exact mechanisms underlying intracellular processing of lipoplexes still remain to be explored.

2.2.1 Transfection experiments

In a typical transfection experiment, the cells are incubated with lipoplexes in transfection medium and the lipoplexes are left to sediment onto the cells, which is a diffusion-limited process. With increasing lipoplex concentration in the transfection medium or longer incubation time, more lipoplexes can sediment and attach to the cells during incubation. On the one hand, this results in a higher number of transfected cells. On the other hand, a higher number of lipoplexes taken up per cell also leads to increased protein expression levels. The fact that the probability of lipoplexes to hit the cell surface and thus the resulting transfection efficiency are directly correlated with the incubation time can be seen in figure 2.3. Here, human hepatocarcinoma cells (Huh 7) and human adenocarcinoma cells (A 549) were incubated at a constant initial dose of mRNA encoding for EGFP and the incubation time was varied. After 25 hours, transfection efficiencies were determined using flow cytometry (FC). The decreasing transfection efficiency for long incubation in the case of Huh7 cells is most probably due to toxic effects caused by the carrier lipids. For most of the experiments described in this thesis, an incubation time of one hour was chosen. Ideally, the kinetics of protein expression after transfection were to be captured right from the start of expression, which is rather immediate for mRNA transfection. However, too short incubation would bring about the disadvantage of significantly lower numbers of transfected cells as can be seen in

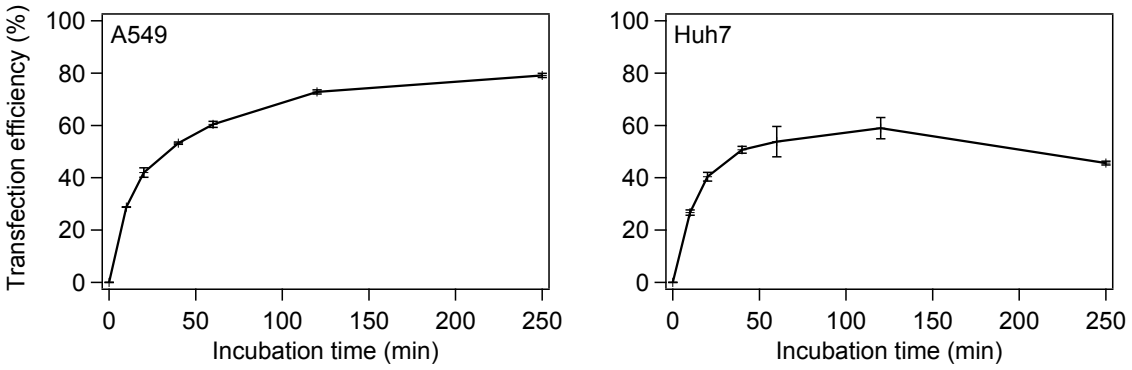

Figure 2.3: *Time-dependent mRNA transfection efficiency measured by flow cytometry in A549 and Huh7 cells at constant initial mRNA dose. Mean transfection efficiencies and standard deviations are shown.*

figure 2.3. Here, the number of transfected cells after increasing incubation time was determined for triplicates (n=5000) using FC.

2.3 Determination of lipoplex size

A package size representing the number of mRNA molecules per lipoplex had to be assumed for the quantitative models of mRNA transfection. To this aim, the diameters of the mRNA lipoplexes were determined using fluorescence correlation spectroscopy (FCS) and dynamic light scattering (DLS).

FCS is a very sensitive, fluorescence-based method that is frequently used to measure particle numbers, diffusion coefficients, hydrodynamic radii, and kinetic rates online, in a small confocal volume, and at concentrations in the nanomolar range [130, 131]. These parameters can be obtained by correlating the fluctuating fluorescence signal inside the detection volume with itself. To measure the hydrodynamic radii of mRNA-lipoplexes with FCS, the mRNA was fluorescently stained with the intercalating dye *Sybr® Gold* before lipoplex formation.

Similar to FCS, the size and the degree of monodispersity of small particles in solution can be determined with DLS by analysing the intensity fluctuations of Rayleigh-scattered light. However, DLS is a label-free technique so that here, no mRNA staining was required.

Table 2.1: *Overview of mean lipoplex diameters as determined by DLS and FCS*

# bases	DLS mean diameter ± std [nm]	FCS mean diameter ± std [nm]
861	85 ± 2	100 ± 11
1000	107 ± 7	114 ± 5
1239	n.a.	120 ± 10

Table 2.1 shows an overview of the lipoplex diameters determined for three mRNA constructs of

different lengths. The values that were obtained using the two techniques are in agreement within experimental error. A detailed description of how the number of mRNA molecules per lipoplex can be deduced from their size is given in the supplementary of associated publication P1 [54].

2.4 The use of green fluorescent protein as a reporter for protein expression

The bioluminescent properties of jellyfish *Aequorea victoria* have first been described by Davenport and Nicol in 1955 [132]. Almost 50 years later, the introduction of GFP and a whole family of GFP-like fluorescent reporter proteins sustainably revolutionized the field of biological sciences. The engineering of fluorescent reporter proteins showing bright and stable emission of light when excited at specific wavelengths paved the way for quantitative studies on spatio-temporal patterns of protein distributions inside and outside living organisms. In 2008, Shimomura, Chalfie, and Tsien were jointly awarded the nobel prize in chemistry by the Royal Swedish Academy of Sciences for substantially contributing to the discovery and development of GFP [1–3].
GFP consists of 238 amino acids which are arranged in a beta barrel structure, with eleven β-sheets forming a cylinder that contains the chromophore, thus protecting it from the surrounding solvent. The chromophore itself, a tripeptide consisting of three amino acids Serine, Tyrosine, and Glycine (Ser65-Tyr66-Gly67), is formed in an autocatalytic process. This post-translational cyclisation is also known as *maturation*. The rate-limiting step of maturation is oxidation.
Nowadays, GFP as well as GFP-derived variants including RNA mimics of GFP as well as photo-activatable GFP and GFP-fusion proteins are a standard for *in vivo* cell biological investigations. All of them act as non-toxic fluorescent reporters inside live cells, some of them even with distinct subcellular localisation [4–6].
A general prerequisite for quantitative measurements using fluorescent reporter proteins is their linear response. Hence, quenching or saturation of the excited chromophores have to be excluded. Also, depending on the time-scale of the measurement, photobleaching effects have to be taken into account. The recombinant EGFP variant used for the underlying work has bright and stable fluorescence properties, its major excitation peak at $488\,nm$ and its emission peak at $507\,nm$, and folds efficiently at 37° [133]. For all measurements, fluorescence intensity decrease due to photobleaching of EGFP was excluded by control experiments.

2.5 Microstructured single-cell arrays

In recent years, various techniques that describe how to create defined microenvironments for cell growth have been published. Approaches include microwell-based arraying of single and multiple cells [134–141] as well as two-dimensional confinement of cell growth with the help of different micropatterning techniques for adhesion molecules [79, 90, 137, 142, 143].
Arraying cells in microenvironments of tunable size and versatile surface functionalisation guaran-

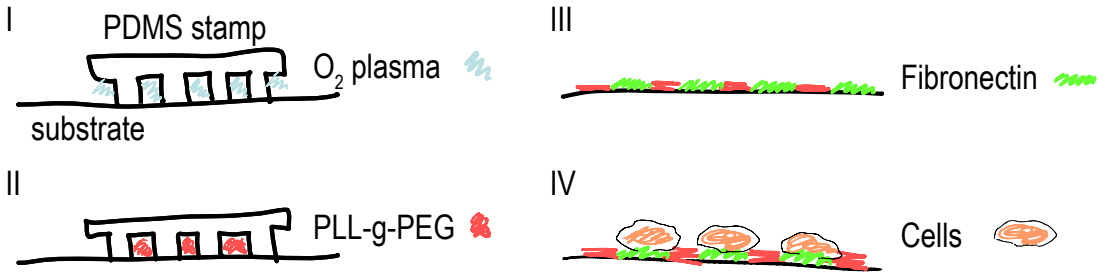

Figure 2.4: *Patterning technique for the creation of single-cell arrays. (I) A polydimethylsiloxane (PDMS) stamp cast from a master produced by photolithography is placed on top of the substrate and treated with oxygen plasma. (II) The parts exposed to plasma are rendered hydrophilic and passivated with PLL-g-PEG. (III) After removal of the stamp, the remaining area is exposed to the extracellular matrix protein fibronectin. (IV) Ideally, seeded cells adhere to the cell-adhesive fibronectin areas only. See appendix C for a detailed protocol.*

tees reproducible and comparable growth conditions for all cells under observation. In addition to that, it facilitates automated image analysis of time-lapse movies, because the investigated cells can no longer migrate due to the confinement to restricted areas of the substrate [84].
For the project presented in chapter 4, the same method as in [90] was used to create single-cell arrays that consist of cell-adhesive, fibronectin-coated squares that are separated by areas functionalised with a cell-repellent block copolymer (PLL-g-PEG). Figure 2.4 shows the manufacturing process of such micropatterned surfaces. Microstructured surfaces were produced by selective oxygen plasma treatment on a plastic substrate with subsequent passivation.

2.6 Quantitative time-lapse fluorescence microscopy

2.6.1 Time-lapse fluorescence microsopy

Fluorescence microscopy is based on the adsorption and subsequent, red-shifted emission of light by both organic and inorganic specimens. As such, it is an essential tool for a broad range of biological, biophysical and biomedical investigations and, more specifically, for time-lapse microscopy studies [84–87, 89, 144, 145].
While the general phenomenon of fluorescence had already been discovered and described in 1852 by Stokes [146], the use of specific fluorochromes to observe biological samples did not begin until the early 20 th century. Today, one can choose from a variety of synthetic and naturally occuring fluorescent dyes and fluorescent proteins [147] for biological imaging.
Figure 2.5 shows the general components of an inverted epifluorescence microscope as it was used for this work. More specifically, a Nikon Eclipse Ti-E microscope equipped with an objective lens (CFI PlanFluor DL-10 x, Phase 1, N.A. 0.30, Nikon) was used. For brightfield imaging, light coming from a halogen lamp is transmitted through the sample. In the fluorescence mode, multispectral light coming from a mercury lamp first passes through a wavelength-selective excitation filter. This light is then reflected by a dichroic mirror and shines on the specimen through the objective. Parts

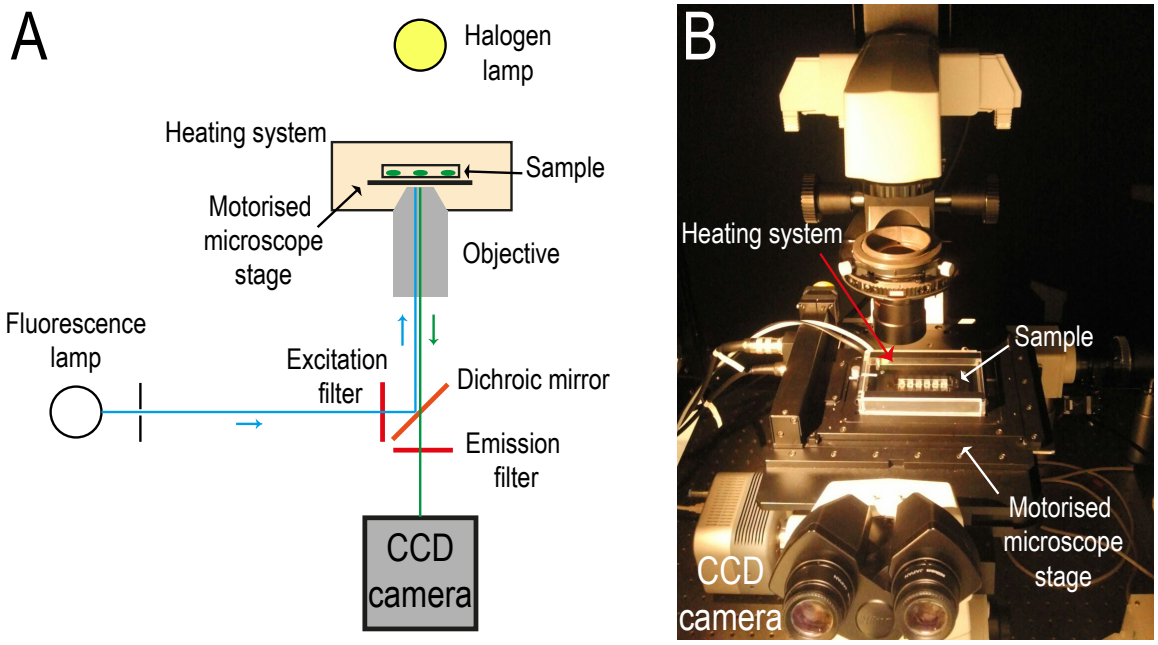

Figure 2.5: *(A) General components of an inverted fluorescence microscope for automated time-lapse microscopy. (B) Photography of the actual setup. The heating system, the sample, and the motorised stage are indicated with arrows.*

of the isotropically emitted fluorescence light are thereafter gathered by the objective and pass through the dichroic mirror. Because the excitation light is generally much brighter than the light emitted by the sample, the emitted light has to be properly separated from the excitation light by an emission filter before reaching the detector, which is a charge-coupled device (CCD) camera in our case. The filter set used for this work was filter set 41024, Chroma Technology Corporation, BP 450-490, FT 510, LP 510-565.

For the multi-position time-lapse microscopy used for this work, two more components that allow extended observation periods under optimal growth conditions for live-cell imaging are added to a basic fluorescence microscope: First, a motorised stage for automated, repeated scanning of multiple positions. Second, a heating chamber that provides optimal growth conditions for biological samples at 37°.

In addition to adverse environmental conditions such as medium depletion or temperature fluctuations, the most significant pitfalls to quantitative live-cell imaging are the inherent focus drift and lamp-intensity fluctuations during long-term imaging. To counter the first problem, our setup profits from a commercial hardware solution for axial focus fluctuations caused by thermal gradients and mechanical instabilities: The so-called *Nikon Perfect Focus System* detects and defines an offset between an axial reference plane and the specimen's focal plane to correct for possible focus drift. The challenge of intensity-fluctuations of the mercury lamp is partly solved by background correction during image analysis (see also 2.7.1).

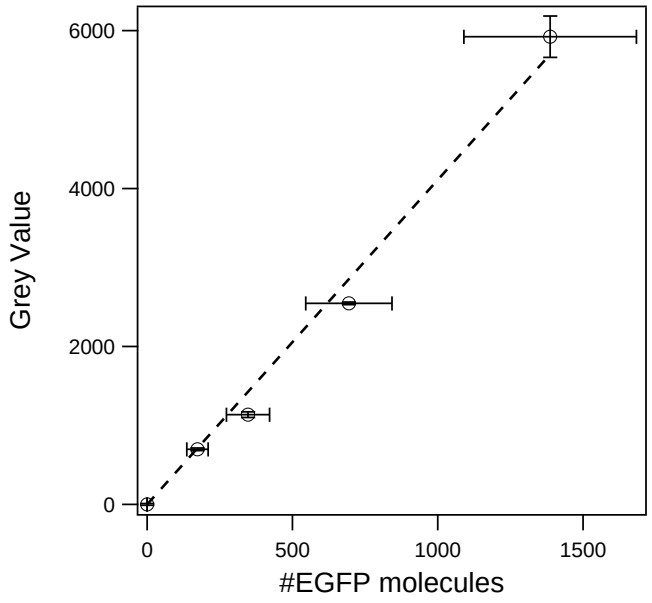

Figure 2.6: *Exemplary calibration curve for one specific set of experimental settings. Microchannels of known dimensions were filled with serial dilutions of an EGFP stock solution of known concentration. This way, pixel intensities were assigned to numbers of fluorescent protein. (For details, see supporting information of P1 [54].)*

2.6.2 EGFP calibration

For all transfection experiments of this thesis, GFP variants (namely EGFP and d2EGFP) were used as fluorescent reporters of protein levels inside the cells. In order to assign actual protein numbers to the intensity values captured by the CCD camera during measurement, we developed a calibration method that is based on microchannels of known dimensions that were filled with EGFP stock solutions of known concentrations. Images of these channels containing serial dilutions of EGFP were taken under the exact same conditions as for the time-lapse measurements. The height of the channels did not exceed the focal depth of the objective used (which is $8.4\,\mu m$ for the 10-fold objective). Thus, by knowing the dimensions of the channels, intensities could be correlated to numbers of fluorescent protein as shown in figure 2.6. This way, we could quantitatively analyse the dynamics of protein expression by conversion of grey values into protein numbers. Our calibration method is described in detail in the supplementary of P1 [54] and the corresponding experimental protocol can be found in the Protocols section.

2.7 Image analysis of time-lapse movies

Given the large data-sets that result from long-term time-lapse measurements, manual image analysis would not only be biased by the person conducting the analysis, but simply not be feasible within a decent amount of time. In a scenario where not only the intensity of the observed objects changes with time, but also the objects of interest are moving at the same time (for example transfected cells producing increasing amounts of protein and migrating on the substrate), automated

image analysis has to reliably track and identify the objects and to detect the intensity of corresponding pixels.
A multi-parameter tracking algorithm based on brightness clustering of pixels that are brighter than a certain threshold value has been presented in [148]. This programm simultaneously tracks moving, fluorescent objects, monitors their cumulative intensity, and can handle events such as dividing or dying cells.
For this work, the *Cell Evaluator*-tracking algorithm as described in [148] and an in-house-development plugin for image analysis, which will be described in chapter 2.7.2, were used to extract data from time-lapse movies.

2.7.1 Background correction

Profound background correction is possibly the most important step in quantitative image processing and crucial for good data quality because incorrect image processing can impair data quality. Image cytometry data can be affected by experimental errors such as intensity distortions caused by the light source or the camera. In addition to the correction of spatial non-uniformity of illumination, systematic intensity fluctuations of the lightsource and autofluorescent effects of cells and cell medium, which are inherent to live-cell imaging, have to be taken into account.
Uneven illumination of the viewfield and pixel-to-pixel variations of the CCD camera's sensitivity are usually overcome by the so-called *flatfield correction*, which eliminates systematic errors so that a uniform signal will result in a uniform output [149]. Subsequent to shading correction, the actual background can be subtracted. Different methods for image normalisation and background correction have been published in the past years [77, 84, 150, 151].
For the underlying work, background correction was implemented in the image-analysis plugins that were used. On the one hand, the *Cell Evaluator* software used for data analysis of the experiments described in chapter 3 possesses a background correction tool that uses either directly measured backgrounds or a background that is created after the measurement. To correct for non-uniform illumination, each pixel was multiplied with a so-called *illumination factor* $f_{illu} = \frac{<b>}{b}$ that accounts for brightness distortions throughout the viewfield by comparing the background for any pixel b and the mean background of the whole image. The resulting pixel values were then given by

$$p_{corrected} = p_{old} \times f_{illu} - < b > . \tag{2.1}$$

More details can be found in the supplementary of [148].
On the other hand, the background correction tool of the plugin for cell-array data analysis detected empty squares, which were then used to calculate a mean background per frame, bc_i, as well as a time-averaged background, $< bc >_t$. The background corrected value $w_{corrected}$ of a square with a cell on it in frame i is given by

$$w_{corrected,i} = (w_{old,i} - < bc >_t) \times \frac{bc_i}{< bc >_t}. \tag{2.2}$$

Squares at the border of the viewfield, i.e. in regions where the shading effect is most pronounced, were automatically deselected. The relative brightness fluctuations within the rest of the analysed image were reasonably small (less than 10 %).

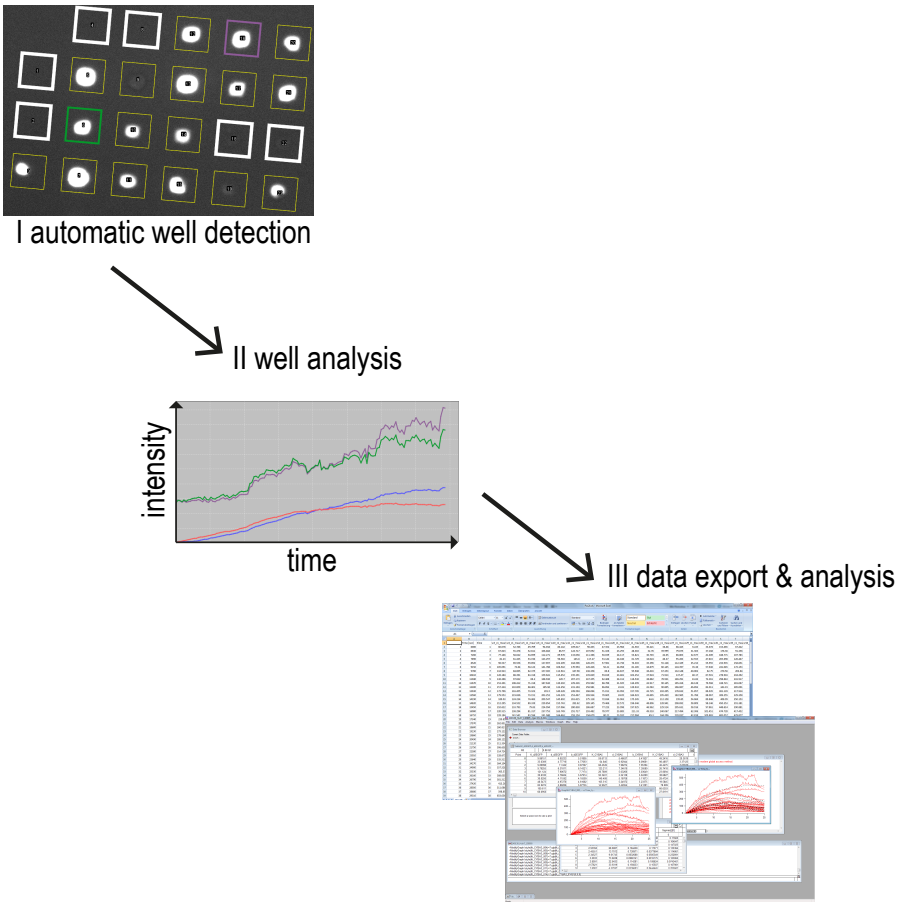

Figure 2.7: *Workflow of automated, single-cell array image analysis and data extraction using the ImageJ plugin. (I) During automatic well detection, the plugin discriminates squares with cells on them (shown in colour) from empty ones (shown in white). (II) During well analysis, pixel intensities of every frame of the analysed image sequence are recorded. Intensity values of empty squares are used for background correction. Here, the green and purple lines show raw data, the blue and red lines are the corresponding time courses after background correction. (III) Subsequently, table-formatted data can be exported for data analysis.*

2.7.2 Automated image analysis of single-cell arrays

Confining cell migration on micropatterned substrates does not only guarantee defined and comparable microenvironments for the observed cells, but also facilitates and speeds up image analysis [84]. To analyse the single-cell protein expression data that are discussed in chapter 4, a versatile plugin for *ImageJ* that was developed in our group was used.

Firstly, the plugin automatically detects squares (called *wells* for historical reasons) that are occupied by cells and those that are empty in the automatic well detection step. The size and

interdistance of these rectangles can be adjusted easily by the user. Then, intensity values of pixels within rectangles are recorded for every frame of an image sequence (well analysis). Empty squares with no cell adhering to them are used for automatic background correction (as described above in 2.7.1). Subsequently, table-formatted data can be exported for further data-analysis, for example using *IgorPro* software. The workflow of the automated single-cell array image analysis is shown in figure 2.7.

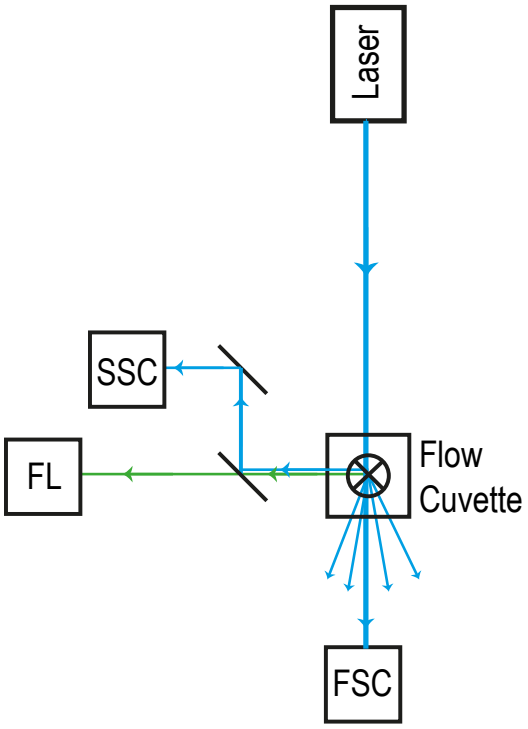

Figure 2.8: *Schematic drawing of flow cytometric setup. A 488 nm laser excites EGFP expressing cells that pass through a flow cuvette. Emitted fluorescent light (FL) is guided to the detector through a long pass dichroic mirror. Forward scattered light (FSC) and sideward scattered light (SSC) are detected, as well.*

2.8 Flow cytometry

Flow cytometry is a high-throughput, laser-based technique that is often used in diagnosis and basic research applications for the assessment of protein expression levels inside cells, to detect binding of labeled antibodies to cell receptors, to measure transfection efficiencies, or for cell sorting and counting. The first impedance-based flow cytometric device was introduced in 1965 [152]. Three years later, the first fluorescence-based flow cytometer was commercially available.
In flow cytometry measurements, the cells are sucked into a flow cuvette, where they are surrounded by the so-called *"sheath fluid"*. Thousands of single cells pass through the cuvette every second, where they are excited by one or more lasers corresponding to the excitation wavelength of the

fluorescent markers to be detected. The following signals are detected: First, the forward scattered light, which is a means to measure the size of a particle passing through the light beam. Second, the sideward scattered light, which is an indicator for the inner structure and hence the granularity of a cell. Third, the fluorescent light emitted by the cell, corresponding to the amount of fluorescent protein produced by it, or indicating binding of labeled antibodies. Depending on the subject of investigation, multicolor setups can be used. Figure 2.8 shows a schematic of the flow cytometric setup that was used for the underlying work.

3 Quantitative analysis of single-cell mRNA and pDNA transfection

Recent advances in the stabilisation of mRNA constructs paved the way for a multitude of mRNA-based studies and gene-therapy applications [46, 48]. Novel mRNA constructs no longer have the disadvantage of low stability and show superior properties over pDNA for the induction of protein expression [153].

In associated publication P1 [54], we discovered another advantage of mRNA transfection, which is its predictability in terms of delivery statistics and protein expression dynamics. In a former study on predictive modeling of non-viral gene transfer, Schwake and co-workers presented quantitative descriptions of the protein expression levels, the transfection efficiencies, and the co-transfection ratio after transfection of cells with two different plasmids at the same time [88]. They reported on simultaneous nuclear entry of complexes, i.e. correlated units comprising several plasmids, and were able to give an estimate of the number of efficiently delivered plasmids per complex.

In the same spirit, we analyzed and compared the timing and the protein expression levels after transfection of eucaryotic cells with either pDNA or mRNA at the single-cell level here. We show a method to determine the distribution functions of mRNA life times. In agreement with the fact that mRNA can be directly translated in the cytosol because, unlike pDNA, it does not have to enter the nucleus in order to be translated, we found significantly shorter onset times of protein expression for mRNA-mediated expression of EGFP.

Moreover, the distribution of the onset times was distinctly narrower for mRNA than for pDNA transfection. In other words, the more complex pDNA transfer process is reflected by a broad distribution of the onset times as well as by higher and less homogeneous protein expression levels as compared to mRNA-induced protein expression. Here, the high copy numbers of each efficiently delivered and subsequently transcribed plasmid, which itself is only a small number, come into effect with their high impact on the resulting number of proteins.

Describing mRNA-mediated protein expression with the analytical solution for a simple biochemical rate model that consists of mRNA-translation and degradation of mRNA and protein, we were able to fit the experimental time courses of mRNA-mediated protein expression. This way, we obtained distributions for the degradation rates and the protein expression rate. The fact that this held true for three different cell types suggests a generic behaviour underlying mRNA-mediated protein expression.

Interestingly enough, a rather simple two-step model for the stochastic delivery of mRNA-lipoplexes correctly describes the dose-response relation. The small number of efficiently delivered mRNA lipoplexes is consistent with previously reported findings for liposome-mediated mRNA transfection where Barreau and coworkers showed that most of the transfected mRNAs are confined in cellular compartments that sequester the molecules away from the cytoplasmatic machineries responsible for mRNA degradation and translation [154]. It had already been reported before that some lipoplexes remain trapped in perinuclear vesicles [155]. The strength of our method to determine mRNA life times lies in the fact that such effects can not bias our results. All degradation rates are assessed on the basis of released mRNA only. As a consequence, the overestimation of life times due to analysis of total mRNA numbers per cell, regardless of the fact that some of them are not even accessible for endogeneous degradation, is circumvented. The multi-level kinetic model that is presented in chapter 5 is an expansion of the model presented in P1.

The experimental data and the resulting conclusions are presented in the following publication P1 [54].

3.1 Associated publication P1

Single-cell mRNA transfection studies: Delivery, kinetics and statistics by numbers

By

Carolin Leonhardt*, Gerlinde Schwake*, Tobias R. Stögbauer, Susanne Rappl, Jan-Timm Kuhr, Thomas S. Ligon, Joachim O. Rädler

published in
Nanomedicine: *Nanotechnology, Biology, and Medicine*, vol.10, issue 4, feature article, 9 pages, 2014.
doi: 10.1016/j.nano.2013.11.008

BASIC SCIENCE

Nanomedicine: Nanotechnology, Biology, and Medicine
10 (2014) 679–688

Feature Article

nanomedjournal.com

Single-cell mRNA transfection studies: Delivery, kinetics and statistics by numbers

Carolin Leonhardt[a,1], Gerlinde Schwake[a,1], Tobias R. Stögbauer, PhD[a], Susanne Rappl[a], Jan-Timm Kuhr, PhD[b], Thomas S. Ligon, PhD[a], Joachim O. Rädler, PhD[a,*]

[a]*Faculty of Physics and Center for NanoScience (CeNS), Ludwig-Maximilians-Universität, München, Germany*
[b]*Institut für theoretische Physik, Technische Universität Berlin, Berlin-Charlottenburg, Germany*

Received 15 March 2013; accepted 18 November 2013

Abstract

In artificial gene delivery, messenger RNA (mRNA) is an attractive alternative to plasmid DNA (pDNA) since it does not require transfer into the cell nucleus. Here we show that, unlike for pDNA transfection, the delivery statistics and dynamics of mRNA-mediated expression are generic and predictable in terms of mathematical modeling. We measured the single-cell expression time-courses and levels of enhanced green fluorescent protein (eGFP) using time-lapse microscopy and flow cytometry (FC). The single-cell analysis provides direct access to the distribution of onset times, life times and expression rates of mRNA and eGFP. We introduce a two-step stochastic delivery model that reproduces the number distribution of successfully delivered and translated mRNA molecules and thereby the dose–response relation. Our results establish a statistical framework for mRNA transfection and as such should advance the development of RNA carriers and small interfering/micro RNA-based drugs.

From the Clinical Editor: This team of authors established a statistical framework for mRNA transfection by using a two-step stochastic delivery model that reproduces the number distribution of successfully delivered and translated mRNA molecules and thereby their dose-response relation. This study establishes a nice connection between theory and experimental planning and will aid the cellular delivery of mRNA molecules.

Key words: mRNA transfection; Non-viral gene delivery; Expression kinetics; Single-cell studies; Pharmacokinetics

Nucleic acid transfer is widely used in basic research as well as biomedical applications. In recent years, novel stabilized mRNA constructs have become more prevalent in therapeutic applications showing superior properties compared to plasmid DNA.[1–3] This progress is mostly due to the discovery of 5′ mRNA anti-reverse cap analogues (ARCA), to the insertion of additional untranslated regions, and to poly(A) tails that significantly promote and prolong efficient translation of foreign mRNA inside cells.[4–10] In general, mRNA delivery has considerable advantages over pDNA delivery in gene therapy applications. Firstly, mRNA does not require transfer into the nucleus and hence mRNA transfection is also effective in non-dividing cells, which is a major drawback of pDNA transfection.[11–13] This makes mRNA a particularly strong therapeutic agent in dendritic cells which are otherwise hard to transfect.[14–16] Secondly, immunogenic response to mRNA activated by Toll-like receptors (specifically TLR3) is less pronounced compared to unmethylated CpG motifs of DNA recognized by TLR9.[2,9,17,18] In addition, mRNA transfection remains transient, preventing the risk of permanently integrating into the genome. Hence, mRNA delivery is of increasing interest for future biomedical applications in particular with regards to strategies that aim to use mRNA as a programmable device for controlled intracellular mRNA targeting and *in situ* logic evaluation of disease-related conditions.[19–24]

The major hurdle to clinical trials remains the delivery of nucleic acid to eukaryotic cells. As a result, an ongoing search is still underway for non-viral delivery methods that are optimized for efficient and controlled delivery of mRNA. Since the first non-viral delivery of mRNA using cationic lipids by Malone, Felgner and Verma,[25] many synthetic delivery systems were found to be effective for mRNA delivery, with generally better efficiency found for liposomes than for polyplexes.[26–31] It is

Financial support by the Elite Network of Bavaria is gratefully acknowledged by CL. This project was supported by the German Excellence Initiative of the Deutsche Forschungsgemeinschaft (DFG) via the Excellence Cluster "Nanosystems Initiative Munich" (NIM), the Sonderforschungsbereich "Nanoagents" SFB 1032, and by the EU-FP7 project ''NanoTransKinetics".

*Corresponding author.
E-mail address: raedler@lmu.de (J.O. Rädler).
[1] These authors contributed equally to the work.

1549-9634/$ – see front matter © 2014 The Authors. Published by Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.nano.2013.11.008

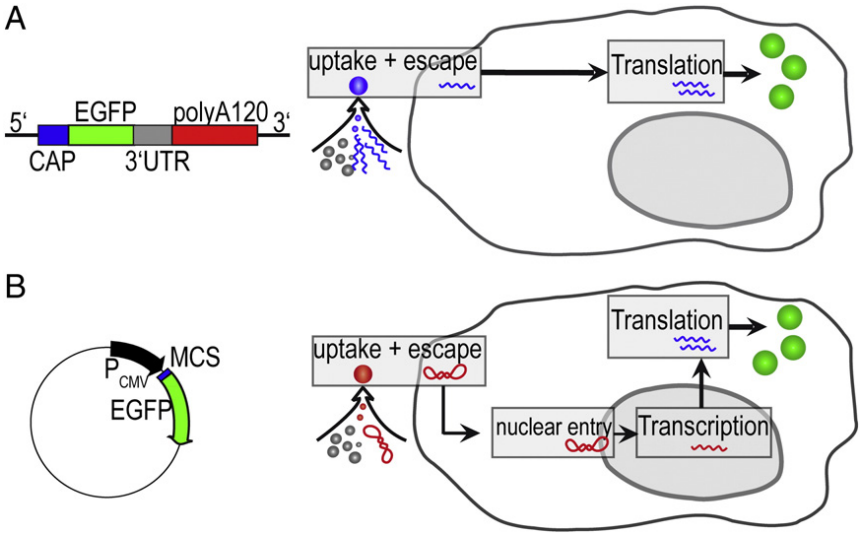

Figure 1. Comparison of mRNA and pDNA Vectors (both gene vectors encoding for the same eGFP protein) and their respective uptake pathways. (**A**) Linearized RNA (1192 bases) furnished with a stabilizing CAP sequence, an enhancing UTR sequence, and poly-(**A**) tail. (**B**) pDNA (4733 base pairs) under the control of the CMV promoter. The vector transfer under identical transfection protocols differs because mRNA is translated after endosomal escape, while plasmid DNA must be transferred into the nucleus for the initiation of transcription.

generally accepted that both mRNA as well as pDNA are translocated via endosomal uptake, cytosolic release and - in case of pDNA - nuclear entry. However, mechanistic insights are mostly limited to assessment of changes in the transfection efficiency as a function of biochemical or structural variations of the carrier. A full pharmacokinetic model, which in principle has been established using compartment models and rate equations,[32–34] lacks validation due to the multitude of kinetic rates. In comparative studies, it was shown that mRNA transfection compared to pDNA transfection is faster and yields a larger fraction of transfected cells.[27,35] However, a more detailed and quantitative understanding in particular of artificial mRNA delivery is of increasing importance for gaining a systems-level description of the kinetics of RNA-based devices.[21,24,36] The degree of predictive power describing synthetic RNA expression level and timing will nevertheless depend on the degree of accuracy with which the transfer efficiency and transfer kinetics can be described. Moreover, predictive modeling of mRNA transfection will be instrumental for the advancement of mRNA based therapies. Yet, any non-viral delivery is inherently stochastic and the expression level and timing of every single cell is different. Hence, measurements at the single-cell level and analysis of the corresponding distribution functions are necessary to acquire the true population response in transfection experiments. Using single-cell analysis, we recently showed that in the case of pDNA transfection, the distribution of gene expression levels can be reproduced using a stochastic model.[37] Similarly, a recent statistical analysis of nanoparticle dosing exhibited Poisson-type distribution in the number of nanoparticles being taken up.[38]

Here, we study gene expression after non-viral delivery of synthetic mRNA analyzing single-cell expression traces in terms of numbers of complexes delivered and numbers of proteins being expressed. Using single-cell fluorescence time-lapse imaging and FC, we monitored expression of a cap-stabilized mRNA vector encoding for eGFP. Single-cell fluorescence time-courses were fitted based on rate equations for translation and mRNA/eGFP degradation yielding the onset time distribution, mRNA/eGFP degradation rates, and the expression rate. The mRNA expression model applies to at least three different cell lines. We interpret the cell-to-cell variability in eGFP levels, i.e. the distribution of expression rates, in terms of number of successfully delivered and translated mRNA. The latter is estimated using a two-step stochastic delivery model. The model assumes delivery of mRNA in finite size complexes that are taken up stochastically by endosomes and randomly released from endosomes into the cytosol. The model quantitatively reproduces the dose–response relation and yields the correct shape of the distribution function. As such, this work represents an advance in predictive modeling of mRNA transfection for quantitative gene expression studies, which we believe will be particularly useful for research on siRNA and miRNA kinetics.

Methods

pDNA and mRNA-vectors

Two different vectors for pDNA and mRNA transfection were designed. The peGFP-N1-Vector (commercially available at BD Biosciences Clontech, Germany, 4733 base pairs) is the standard eGFP vector. As an mRNA reference construct for *in vitro* transcription, we designed a vector that is based on the pSTI-A120-vector (4746 base pairs, transcript 1192 bases), which has previously been described in literature.[10] The complete vector map is presented in Figure S1. Both vectors contain the same eGFP gene but differ in their promoter region: The peGFP-N1-Vector has a strong CMV-promoter for expression *in vitro*. The mRNA is generated with a commercial *in vitro* transcription kit from the pSTI-A120-vector under the control of the T7 promoter. The backbone of both vectors is

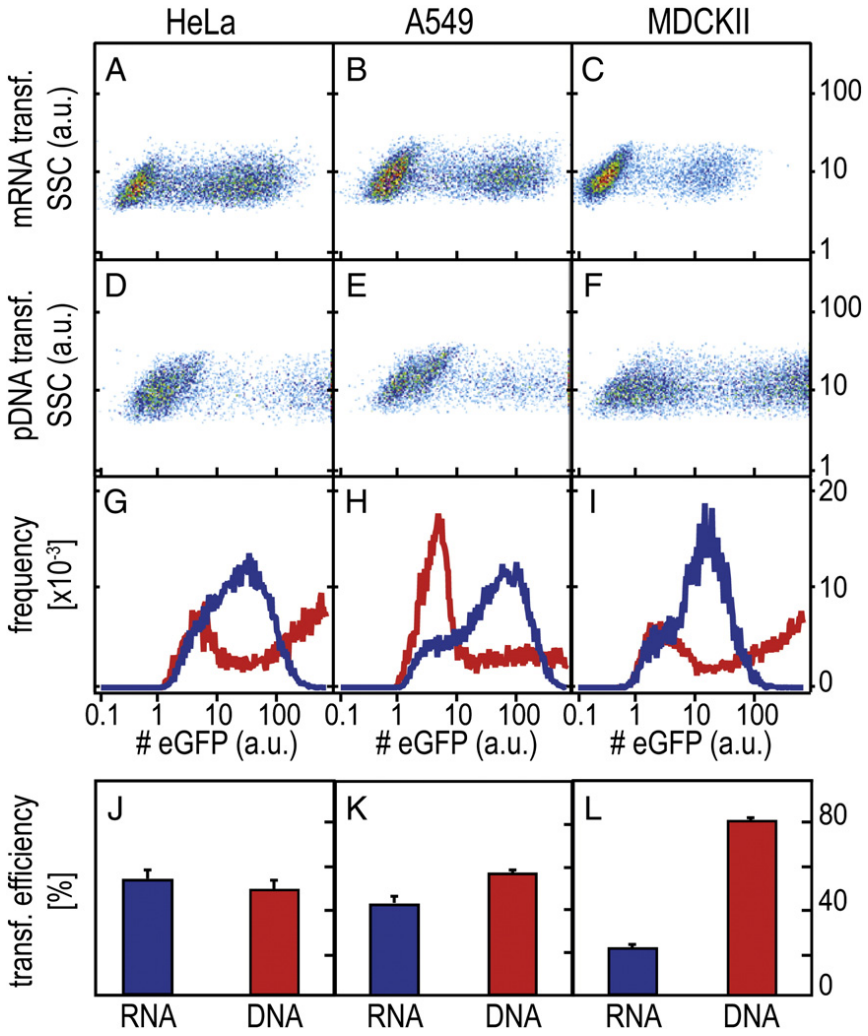

Figure 2. Representative FC scatter plots for mRNA- and pDNA- mediated eGFP expression in three different cell lines (arbitrary units). (**A**-**F**) Two-dimensional scatterplots (sideward scatter vs. fluorescence intensity) for HeLa, A549 and MDCKII cells 25 h post-transfection with mRNA and pDNA. (**G**-**I**) Average fluorescence intensity per fluorescent cell (RNA data are shown in blue, DNA data are shown in red); (**J**-**L**) Percentage of fluorescent cells (mean ± SD).

based on the pCMV-Script vector. pSTI-A120 has a 120-bp poly(A) tail and a 3′ untranslated region (UTR) from human β-globin enabling *in vitro* transcription of polyadenylated RNA.

To generate *in vitro*-transcribed mRNA (IVT RNA), the plasmid is linearized downstream of the poly(A) tract by SapI digestion and purified by phenol/chloroform extraction and sodium acetate precipitation. One μg of the linearized vector is used as a template for the *in vitro* transcription reaction using the Biozym Kit (MessageMAX™ T7 ARCA-Capped Message Transcription Kit). Having an Anti-Reverse Cap Analog (ARCA) ($m_2{}^{7,3'\text{-}O}$G[5′]ppp[5′]G) cap on the 5′ end, ARCA cannot be incorporated in the reverse orientation. Thus, 100% of the caps in the produced IVT RNA are in the correct orientation, increasing the translation efficiency of the IVT RNA.[10,39]

Materials

FBS, Leibovitz's L-15 Medium (Gibco), Lipofectamine™2000, OptiMEM (Gibco) and Sybr Gold were purchased from Invitrogen, Germany. Syto RNAselect was purchased from Life Technologies, Germany. 6-well culture plates (Falcon) were purchased from VWR International GmbH, Germany. Sterile PBS was prepared in-house. Ham's F-12K, MEM, DMEM and Trypsin-EDTA were purchased from c.c.pro GmbH, Germany.

Cell culture

A human alveolar adenocarcinoma cell line (A549, ATCC CCL-185) was grown in Ham's F12K medium supplemented with 10% FBS. HeLa cells (ATCC CCL-2) were cultured using minimum essential medium (MEM) with Earle's salts and L-Glutamine supplemented with 10% fetal bovine serum (FBS). A Madin-Darby Canine Kidney epithelial cell line (MDCKII, ATCC CCL-34) was cultured in DMEM with 4,5 g/L glucose and 110 mg/L pyruvate, supplemented with 10% fetal bovine serum. All cell lines were grown in a humidified atmosphere at 5% CO_2 level.

Transfection

The cells were transfected with equimolar amounts of pDNA and mRNA for FC measurements and with equal weight amounts

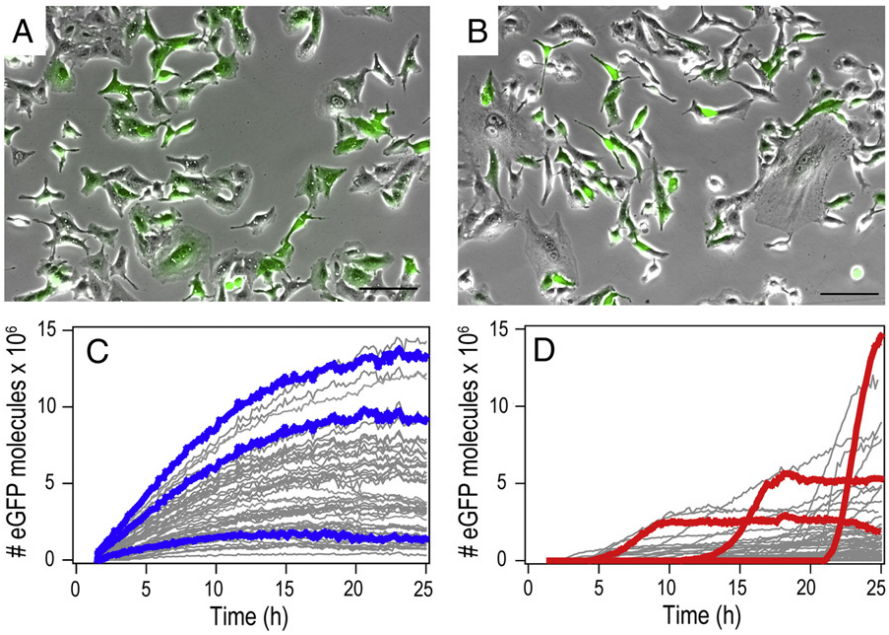

Figure 3. mRNA- and pDNA-mediated gene expression kinetics. (**A**, **B**) Exemplary images of an average transfection of A549 cells 25 h post-transfection (overlay of bright field and eGFP fluorescence image. Scale bars 100 μm). (**C**, **D**) Representative fluorescence time-courses of eGFP gene expression after transfection with mRNA (**C**) and pDNA (**D**). To highlight the characteristic differences, we chose and color-labeled three exemplary time-courses each. mRNA expression shows early onset and continuous rise in the eGFP level, while pDNA expression exhibits delayed onsets and S-shape expression time-courses.

of pDNA and mRNA for single-cell measurements (see Supplementary). The same transfection reagent (Lipofectamine2000®) and the same standard transfection protocols were used for pDNA and mRNA delivery. For transfection with fluorescently labelled mRNA, we followed the standard protocols for labelling mRNA with Sybr Gold/Syto RNAselect and prepared lipoplexes with labelled mRNA.

Data acquisition and quantitative image analysis

Live-cell imaging was performed on a motorized inverted microscope (Nikon, Eclipse Ti-E) equipped with an objective lens (CFI PlanFluor DL-10×, Phase1, N.A. 0.30; Nikon) and with a temperature-controlled mounting frame for the microscope stage. To acquire cell images, we used a cooled CCD camera (CLARA-E, Andor). A mercury light source (C-HGFIE Intensilight, Nikon) was used for illumination and a filter cube with the filter set 41024 (Chroma Technology Corp., BP450-490, FT510, LP510-565) was used for eGFP detection. An illumination shutter control was used to prevent bleaching. Images were taken at 10 fold magnification with a constant exposure time of 1300 ms at 10-minute intervals for at least 25 hours post-transfection. Fluorescence images were consolidated into single-image sequence files. Negative control images were taken to assess lamp threshold values and were subtracted from corresponding image sequence files to eliminate autofluorescence effects. Using SINGLECELLTRACKER, an in-house-development software based on ImageJ,[40] fluorescence intensities were integrated over cell contours and corrected for background noise. The software calculates the cells' fluorescence over the entire sequence and connects corresponding intensities to time-courses of the fluorescence per cell.

eGFP quantification and calibration

To calculate numbers of eGFP molecules from grey values of the recorded time-lapse movies, a calibration-channel system was developed. Micro channels of known dimensions were filled with eGFP solutions of defined concentrations. Images of the channels were taken under the same experimental conditions as the monitored expression kinetics data, corrected for background and analysed to get calibration curves. For a detailed description of the calibration method, see Supplementary.

Flow cytometry

eGFP fluorescence intensity in cells was measured by FC (Partec, CyFlow space). Flow cytometer settings were adjusted to discriminate transfected and non-transfected cells. The Windows™ FloMax® software package was used for data analysis. See Supplementary for additional information.

Results

mRNA vs. pDNA transfection

In a first set of experiments, mRNA-mediated transfection was quantified using FC and compared to pDNA-mediated transfection as a reference. As schematically depicted in Figure 1, the design of the mRNA vector (Figure 1, *A*) was chosen for maximal analogy to the pDNA vector. The pDNA vector is a commercial eGFP plasmid equipped with a CMV promoter (Figure 1, *B*). The mRNA construct consists of polyadenylated RNAs enabling *in vitro* transcription under the control of the T7-promoter and contains 2 sequential human β-globin 3′UTRs as well as the anti-reverse cap analog (ARCA) (see also Methods,

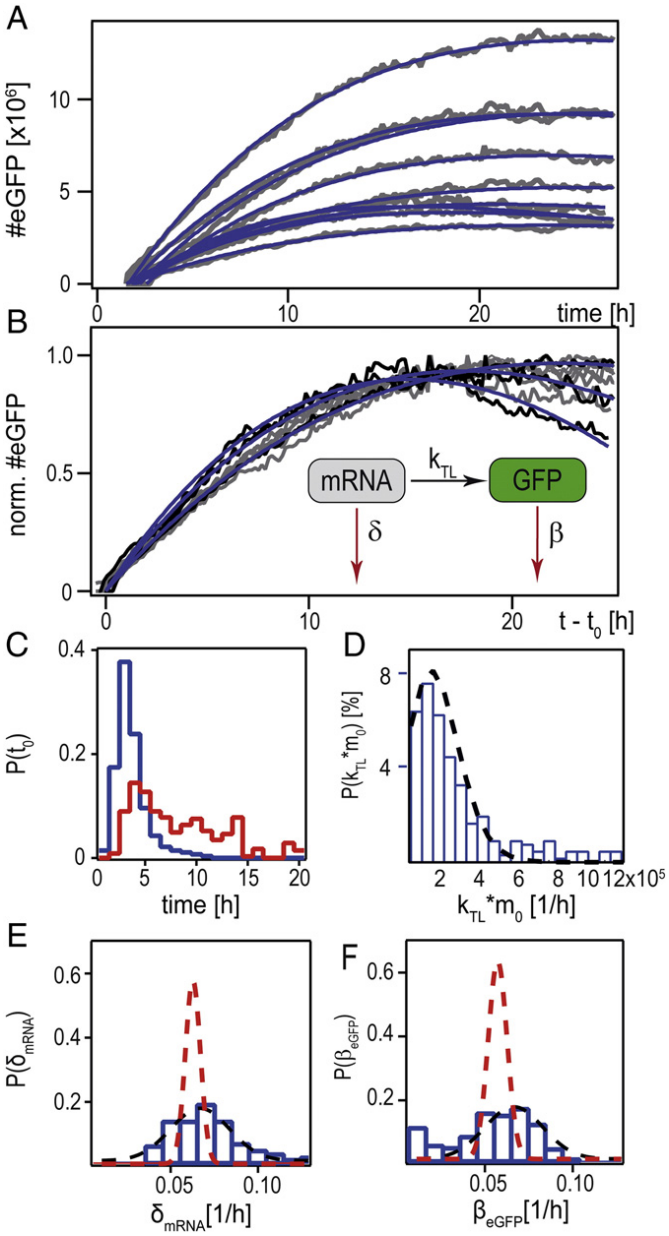

Figure 4. Single-cell mRNA translation, analyzed by a kinetic rate model. (**A**) Time-courses of eGFP expression after mRNA transfection (gray lines). Blue lines are fits according to the rate equation model (shown schematically as insert in (**B**)). (**B**) Shows the same data as (**A**), normalized to their maximal value and shifted by their fitted onset times, t_0. (**C**) Distribution of the onset time t_0 (mRNA data shown in blue, pDNA data shown in red). (**D**) Distribution of the expression rate $k_{TL} \cdot m_0$. (**E**) Distribution of the mRNA degradation rate. The black dashed line shows the Gaussian fit to the experimental data, whereas the red dashed line is the Gaussian fit to simulated data (see Supplementary) (**F**) Distribution of the eGFP decay rate. Dotted lines represent the Gaussian fit to experimental (black) and simulated (red) data.

Supplementary).[10] To collate the outcome of the transfection experiments, identical transfection protocols were followed for mRNA and pDNA transfection using the commercial cationic lipid agent Lipofectamine2000®.

The FC data shown in Figure 2 were taken 25 hours post-transfection. The scatterplots with the fluorescence intensity on the x-axis and the sideward scattering signal on the y-axis show consistent bimodal populations. Both mRNA and pDNA mediated transfection exhibit eGFP-expressing cells and cells that do not express any eGFP. However, for three different cell types, the fluorescence level of eGFP expressing cells in case of pDNA mediated expression is more broadly distributed and shifted towards higher values than the eGFP distribution appearing in mRNA transfection. This effect is also seen in the integrated representation, where the distribution of the average number of eGFP molecules per eGFP expressing cell is shown (Figure 2, *G–I*). Here, pDNA transfection is shown in red and mRNA transfection in blue. Note that for pDNA transfection, 22% (HeLa), 7% (A549), and 28% (MDCKII) of the cells exhibit eGFP expression levels of 1000 (a.u.) and higher that are not shown for better clarity. In the last row (Figure 2, *J-L*), the percentage of transfected cells are depicted, which is a direct measure of the transfection efficiencies. We find slightly lower percentages of transfected cells for mRNA-transfected cells compared to pDNA-transfected cells except for MDCKII cells, which feature higher transfection for pDNA vectors.

Single-cell mRNA expression kinetics

The most revealing difference between transfection with mRNA and pDNA is seen in the single-cell expression kinetics retrieved from time-lapse studies (Figure 3). Typically, beginning after 1.5 hours of incubation, fluorescence microscopy movies were taken over 25 hours using automated time-lapse microscopy. The total fluorescence intensity of each single cell was followed by image analysis[40] and converted into the number of eGFP molecules per cell (see Supplementary). Figure 3 shows two typical microscopy images of transfected cells 25 hours post-transfection (Figure 3, *A* and *B*). Bright field and fluorescence images were overlaid to illustrate the fraction of transfected cells. Figure 3, *C* and *D* show gene expression time-courses of single cells. To highlight the characteristic differences in the expression kinetics, we picked three representative traces each and show them in color. While mRNA-transfected cells show an early and steady rise to a maximum with a subsequent decrease, pDNA transfection results in sigmoidal intensity time-courses with a steady-state level of eGFP expression and random onset times. In contrast to the ubiquitous early onset of eGFP expression with mRNA that mainly occurs within 5 hours after transfection, the onset of eGFP expression after transfection with pDNA is spread over the range of 2 hours to 20 hours.

Modeling mRNA expression

Since mRNA transgene expression solely involves translation, quantitative modeling reduces to a simple biochemical reaction scheme defined by three kinetic rates as shown in Figure 4, *B*. The schematic shows a rate equation model for mRNA expression consisting of translation, mRNA, and eGFP-degradation. The model is described by the following set of equations for the changes in the number of eGFP molecules, *G(t)*, and the number of mRNA molecules, *m(t)*:

$$\frac{d}{dt}G = k_{TL} \cdot m - \beta \cdot G \qquad (1)$$

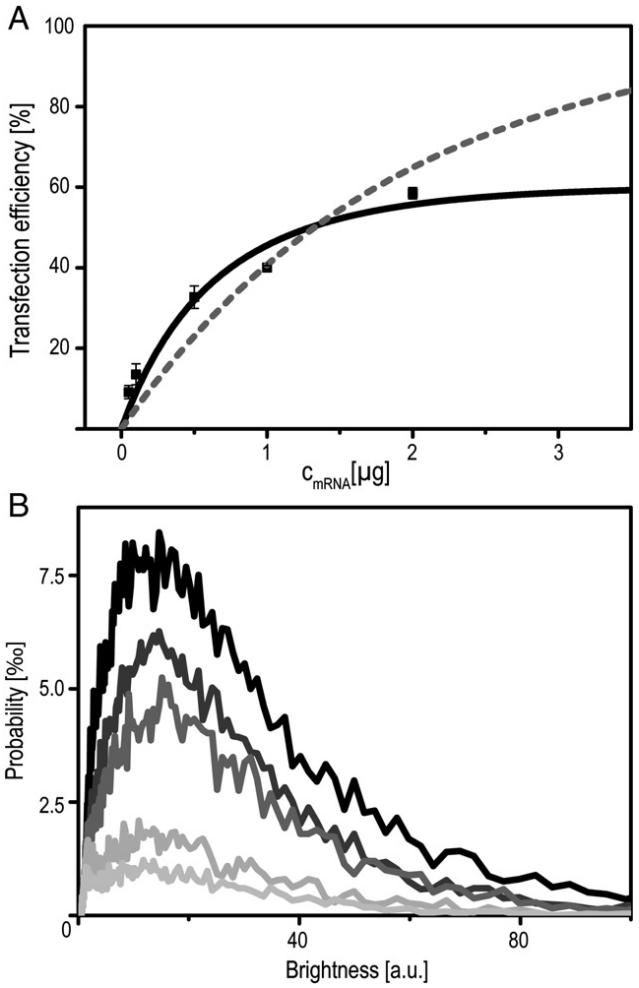

Figure 5. Dose–response relation. (**A**) Percentage of positively transfected A549 cells as a function of increasing amount of mRNA (0.05/0.1/0.5/1/2 µg). Squares correspond to FC data. The dashed grey line is a single-Poisson fit, the black line is a double-Poisson fit according to our stochastic delivery model. (**B**) Corresponding fluorescence intensity distributions as measured by FC (bottom to top with increasing mRNA dose).

$$\frac{d}{dt}m = -\delta \cdot m \tag{2}$$

where k_{TL} denotes the translation rate and δ and β the degradation rates of mRNA and eGFP, respectively. With t_0 being the time of expression onset and the initial conditions $G\ (t_0) = 0$ and $m\ (t_0) = m_0$, the following solution for the number of eGFP molecules is obtained:

$$G_{mRNA}(t) = \frac{k_{TL} \cdot m_0}{\delta - \beta} \cdot \left(1 - e^{-(\delta-\beta)(t-t_0)}\right) \cdot e^{-\beta(t-t_0)} \tag{3}$$

Applying Eq. 3 to the experimental time-courses, the data are indeed well fitted. The blue curves in Figure 4, *A* show exemplary best fits to single-cell time-courses (from a total of 281 time-courses). There are four free parameters: the onset time t_0, the product of translation (k_{TL}) and initial number of effectively translated mRNA molecules (m_0), as well as mRNA and protein degradation rates (δ and β). Eq. 3 entails a time-course showing an exponential increase with rate δ-β and a long-term decay with decay rate β (see Supplementary). Each fit yields an individual set of parameters. Figure 4, *C-F* presents the corresponding distribution of the best-fit parameters, which will be discussed in the following.

Expression onset time distribution

In Figure 4, *C*, the onset time of mRNA (blue) is shown in comparison to the onset time for pDNA transfection (see Supplementary). The faster transfer of mRNA is clearly documented in this distribution. In the case of A549 cells shown here, the onset time distribution after transfection with mRNA peaks approximately 3 hours after transfection and hardly shows any delayed expression onset events after 5 hours, whereas the pDNA onset time distribution is spread over the interval between 2 and 20 hours post-transfection. The time-distribution is an indirect, yet quantitative measure for the transfer time of delivery. As known from microscopy studies, endosomal uptake already starts 10–30 minutes after transfection.[30,41] Therefore, the measured delay in case of mRNA transfer must be limited by endosomal escape rates. Remarkably, mRNA expression onset ceases after 10 hours, indicating that no more endosomes lyse or (more likely) that mRNA molecules are degraded in acidic late endosomes. The broadly distributed onset times for pDNA are associated with rare nuclear entry events, which are believed to occur predominately during mitosis.

mRNA degradation rates

Figure 4, *E* shows the distribution of the mRNA degradation rate retrieved from fitting single-cell time-courses with the described model. The average mRNA degradation rate of 0.062/h (corresponding to an mRNA life time of $t_{1/2} \approx 11$ hours) is in rough agreement with the literature value of 0.028/h.[10] The value is clearly smaller than the degradation rate of endogenous mRNA ($\delta < 0.14$/h[42]), which is consistent with the reportedly higher stability of ARCA capped mRNA vectors. The distribution of mRNA degradation is well described by a Gaussian with half-width 0.024/h. This variability in the degradation rate is on the order of the so-called "extrinsic noise" in eukaryotes.[43] The values for the degradation of eGFP (with a mean of 0.056/h) are higher than values that have been reported previously.[42,44] In general, it is noteworthy that the single-cell analysis yields estimates for δ and β with high accuracy. The Gaussian fit yields mean values with less than 6% relative error. Knowing the degradation rates is of great value for the improvement of novel vectors and capping sequences. Furthermore, the degradation times are a key to predicting the time-course of expression. In fact, analysis of Eq. (3) predicts that the maximum of expression is reached approximately at $t_{max} = 17$ h. The time point of half maximum expression value in the declining late phase of expression is $t_{1/2} = 45$ h. The latter is important because it is a measure for the duration of the transient mRNA expression. Note that Eq. 3 also holds for the case $\delta < \beta$ (see Supplementary). Moreover, the expression rate $k_{TL} \cdot m_0$ and the difference in the degradation rates (δ-β) both determining the amplitude and hence the maximal expression levels, are uncorrelated (see

Supplementary, Figure S3C). In Figure 4, *E* and *F*, Gaussian fits to simulated data are additionally shown. For simulation, we used the experimentally measured mean degradation rates (see Supplementary). These fits should represent intrinsic noise only, which accounts for about 30% of the total noise. The additional width of the experimental data can be attributed to extrinsic sources of noise involved in the gene transfer process. The kinetics of mRNA proves to be generic because different cell types show the same mRNA expression curves (see Supplementary).

A stochastic delivery model by Numbers

It is generally understood that mRNA as well as pDNA delivery via artificial, non-viral vectors is stochastic and dominated by rare processes. In the case of mRNA transfection, the limiting steps are endosomal uptake, endosomal lysis, and mRNA release from lipoplexes. Here, we ask the question whether the measured distribution of expression levels can be reproduced in a stochastic rate model, where each step is assumed to be described by a random process with defined transition probability. The fact that a large fraction of cells does not express eGFP at all indicates that there is a finite probability that no nucleic acid is successfully transferred. Figure 5, *A* shows the dose–response curve in terms of the percentage of transfected cells versus the concentration of mRNA in μg RNA per ml transfection medium. The corresponding distribution of eGFP expression levels can be seen in Figure 5, *B*. Data were taken 25 h after transfection using FC. The number of transfected cells monotonically increases with mRNA dosage. It is instructive to describe the transfection process in terms of number of lipoplexes: Lipoplexes form when cationic lipid liposomes are complexed with nucleic acid. Each lipoplex contains a large average number of mRNA molecules (as discussed below). Hence, the delivery of a single lipoplex results in a burst of eGFP expression. If lipoplexes were delivered by overcoming a single barrier, the dose–response function would be described by a Poisson-like process as represented by the dashed line in Figure 5, *A* (see Supplementary). In this case, the average number of effectively delivered lipoplexes would be $\langle C\rangle_{SP} = 0.5$. However, as shown in Figure 5, *A*, the fraction of transfected cells can be more closely described by a chain of two successive Poisson processes. In this case, the response does not rise up to 100% at large mRNA concentration, which is due to the fact that the two Poisson processes are sequential. A physical interpretation of such a chain of events is shown in Figure 6, *A*: The scheme shows endosomal uptake of lipoplexes, endosomal lysis, and mRNA release from lipoplexes. It is assumed that N endosomes are stochastically loaded with a small number of lipoplexes, L_{eff}, and that subsequently a small fraction of endosomes, N_{eff}, undergoes lysis. These two stochastic steps are modeled as Poisson processes and determine the number of delivered lipoplexes, C. If we assume the lipoplex load L_{eff} to be proportional to the mRNA concentration, i.e. $L_{eff} = \lambda \cdot c_{mRNA}$, we obtain a two-parameter expression for the dose–response function (see Figure 5, *A* and Supplementary). The best fit yields $N_{eff} = 0.9$ and $\lambda = 1.1\ \mu g^{-1}$, meaning that at the highest dose of 2 μg, an effective number of $L_{eff} = 2.2$ lipoplexes are contained per endosome and that an average of $\langle C\rangle = N_{eff} \cdot L_{eff} = 2$ successfully delivered complexes is obtained. To demonstrate that such a surprisingly small number of effectively delivered lipoplexes is realistic, we assessed the average number of lipoplexes resting on a single cell in an experiment. At a dose of 1 μg mRNA and after one hour

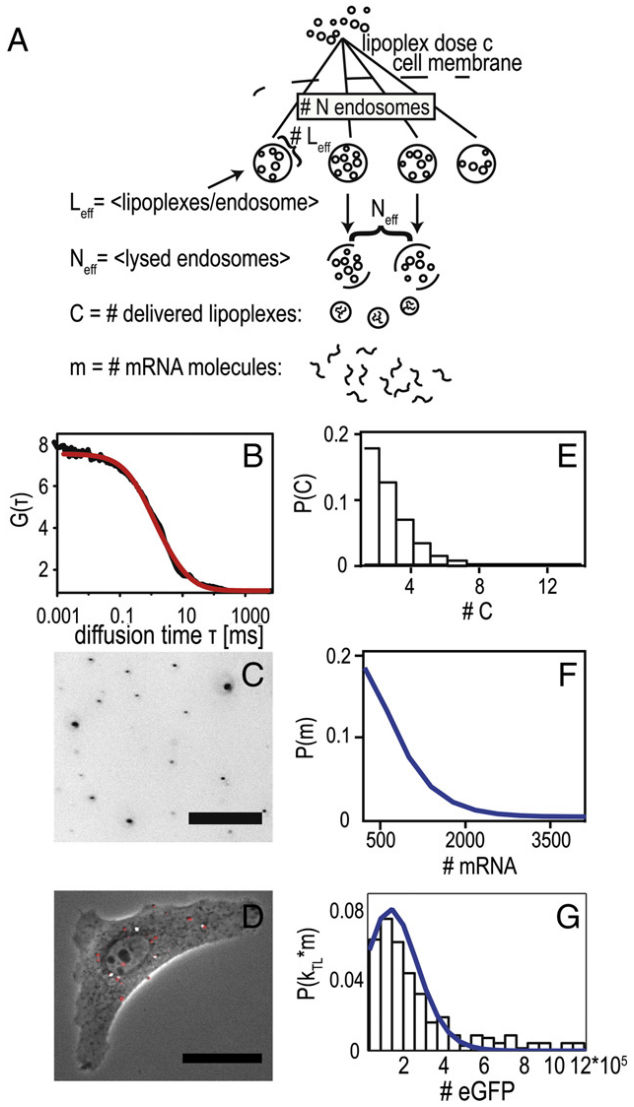

Figure 6. Two-step stochastic mRNA delivery model. (**A**) Schematic drawing of the stochastic uptake of lipoplexes by endosomes, lysis of the endosomes, and release of the mRNA load by lipoplexes. The model reproduces the dose–response relation shown in Figure 6, *A*. (**B**) Fluorescence autocorrelation function of lipoplexes showing an average hydrodynamic radius of $R_{hydr.} = 60$ nm. (**C**) Fluorescence image of fluorescently labeled mRNA lipoplexes adsorbed to a petri dish at the concentration that was used for time-lapse transfection experiments (dose: 1 μg/ml mRNA). Image analysis led to a typical lipoplex density of order 4000/mm^2 corresponding to about 4–8 lipoplexes per cell (intensity scale inverted for clarity, scale bar 25 μm). (**D**) Typical A549 cell five hours after transfection with fluorescently labeled mRNA-lipoplexes (shown in red, scale bar 25 μm). (**E**) Predicted distribution of delivered lipoplexes derived from the dose–response relation. (**F**) Predicted distribution of delivered mRNA molecules, based on an average of 350 mRNA molecules per lipoplex. (**G**) Experimental probability distribution of expression rates ($k_{TL} \cdot m_0$, black bars) derived from single-cell data. Blue line indicates best fit of mRNA distribution to the expression distribution, yielding an approximate translation rate of $k_{TL} = 170$/h.

incubation time, we found a lipoplex surface density of about 4000/mm^2, corresponding to an average of 4–8 lipoplexes per cell (Figure 6, *C*). This number is strongly dependent on incubation time due to the diffusion limited transport of the lipoplexes. After five hours of incubation, the number of lipoplexes doubles as seen in Figure 6, *D*. We can safely assume that almost all lipoplexes that hit the cell surface will be taken up by endocytosis over time as reported by others.[29,30] However, not every endosome releases its lipoplex cargo into the cytosol. We find that a lysis rate of about 25–50% leads to accordance of the experimental dose–response relation with the above theoretical estimate.

A single lipoplex contains an average of $\langle m \rangle = 350$ mRNA molecules. This number is derived knowing the size and packing density of lipoplexes (see Supplementary). The mRNA lipoplexes used here exhibit an average hydrodynamic radius of 60 nm as measured by fluorescence correlation spectroscopy (FCS) (Figure 6, *B*). The structure and packing density have been measured previously using small angle X-ray scattering (SAXS).[45,46] Figure 6, *E* shows the theoretical distribution of delivered lipoplexes based on the double-Poisson model and the mRNA dose that was used for these experiments (1 μg). If this distribution is multiplied with the number of mRNA molecules per lipoplex, we obtain the theoretical distribution of mRNA per cell as shown in Figure 6, *F*. It is noteworthy that the theoretical distribution (Figure 6, *G*, blue curve) is in very satisfying agreement with the shape of the experimental distribution (Figure 6, *G*, black bars) of expression rates. Comparing the theoretical mRNA distribution with the actually measured distribution of expression rates, $k_{TL} \cdot m_0$, we find $k_{TL} = 170$/h. This translation rate, which emerges from the analysis of single-cell expression rates, is in the range of independently published values of translation rates.[42]

Discussion

We studied the expression kinetics of eGFP following transfection mediated by mRNA and pDNA. While pDNA complexes have to enter the nucleus, mRNA molecules released from mRNA lipoplexes can be translated immediately after endosomal escape. Consequently, mRNA-induced expression is profoundly earlier and more homogeneously timed than pDNA-induced expression. This behavior is generic and similar onset time distributions are observed e.g. for HeLa and MDCKII cells (data not shown). The high transfection efficiencies for pDNA transfected cells as compared to mRNA transfected cells might be a result of size-dependent lipoplex uptake that has been reported previously.[47] We determined the pDNA-lipoplexes to be about 230 nm in diameter (data not shown), as opposed to 120 nm for mRNA-lipoplexes. The narrow timing of mRNA expression onset at approximately 3 hours post-transfection is in agreement with the observed timing found for endosomal uptake and release in single-particle tracking studies.[30,41] Therefore, the mRNA expression onset distribution might serve as a valuable indicator for the endosomal release time distribution and could be useful for the advancement of artificial endosomolytic agents. Furthermore, our data imply that mRNA expression modeling can predict the transient course of therapeutic efficacy of mRNA therapeutics in preclinical studies. For example, the development of improved capping sequences of mRNA vectors can be carried out using destabilized eGFP variants. In this case, the protein level decreases substantially faster and long observation times causing experimental difficulties can be circumvented (see Figure S7, Supplementary). Based on kinetic rates obtained in such studies, the time-course of arbitrary gene products with longer half-life times can be inferred. In this context, it should be noted that the half-life of about 12 hours for eGFP determined from single-cell tracks is shorter than previously reported in ensemble measurements, which necessarily average over the somewhat heterogeneous timing of whole populations.[42,44] We also showed that the cell-to-cell variability in the expression levels is well described by a two-step Poisson process. The two-step stochastic model is capable of reproducing the measured dose–response curve consistently with the statistical distribution of expression rates. However, it is limited to transfection *in vitro* and provides only an approximate description of the underlying delivery cascade. The most important element provided by our model is the account of quantal delivery of mRNA in form of lipoplexes, which is in quantitative agreement with the measured distribution functions. The small number of successfully delivered lipoplexes per cell is the key to understanding the stochastic outcome of transfection experiments that inherently allow a finite number of non-transfected cells. More refined modeling has to be done to picture the dynamics of transfection and to reproduce the onset time distribution. Here, computational representation of size-dependent uptake rates, the nature of endosome lysis, and intracellular diffusion need to be solved. Furthermore, computational modeling of extracellular delivery, mimicking *in vivo* situations, needs to be advanced to gain impact on translational medicine.

In our experiments, the single-cell time-courses of mRNA-mediated transfection showed excellent agreement with the standard biochemical rate model of translation. Hence, single-cell analysis enables direct determination of expression rates as well as decay rates for both mRNA and eGFP with great accuracy and provides a quantitative foundation for kinetic studies on mRNA translational regulation as for example RNA interference. The fact that mRNA transfection exhibits a narrow time window of delivery is beneficial for kinetic studies. This advantage should be of practical importance for future time-resolved studies on siRNA knockdown and RNA constructs for programmed gene regulatory operations.

Acknowledgments

We thank Carsten Rudolph for the friendly gift of the vector pSTI-A120, Svenja Lippok for FCS measurements, David Smith for proof-reading of the manuscript, and Maria P. Dobay for helpful discussions.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.nano.2013.11.008.

References

1. Tavernier G, Andries O, Demeester J, Sanders NN, De Smedt SC, Rejman J. mRNA as gene therapeutic: how to control protein expression. *J Control Release* 2011;**150**(3):238-47.
2. Yamamoto A, Kormann M, Rosenecker J, Rudolph C. Current prospects for mRNA gene delivery. *Eur J Pharm Biopharm* 2009; **71**(3):484-9.
3. Andries O, De Filette M, Rejman J, De Smedt SC, Demeester J, Van Poucke M, et al. Comparison of the gene transfer efficiency of mRNA/GL67 and pDNA/GL67 complexes in respiratory cells. *Mol Pharm* 2012;**9**(8):2136-45.
4. Stepinski J, Waddell C, Stolarski R, Darzynkiewicz E, Rhoads RE. Synthesis and properties of mRNAs containing the novel "anti-reverse" cap analogs 7-methyl(3′-O-methyl)GpppG and 7-methyl (3′-deoxy) GpppG. *RNA (New York, NY)* 2001;**7**(10):1486-95.
5. Jemielity J, Stepinski J, Jaremko M, Haber D, Stolarski R, Rhoads RE, et al. Synthesis of novel mRNA 5¢ cap-analogues: dinucleoside P1, P3-Tri-, P1, P4-Tetra-, and P1, P5-Pentaphosphates. *Nucleosides Nucleotides Nucleic Acids* 2003;**22**(5–8):691-4.
6. Zohra FT, Chowdhury EH, Tada S, Hoshiba T, Akaike T. Effective delivery with enhanced translational activity synergistically accelerates mRNA-based transfection. *Biochem Biophys Res Commun* 2007;**358**(1): 373-8.
7. Kuhn AN, Diken M, Kreiter S, Selmi A, Kowalska J, Jemielity J, et al. Phosphorothioate cap analogs increase stability and translational efficiency of RNA vaccines in immature dendritic cells and induce superior immune responses in vivo. *Gene Ther* 2010;**17**(8):961-71.
8. Pesole G, Grillo G, Larizza A, Liuni S. The untranslated regions of eukaryotic mRNAs: structure, function, evolution and bioinformatic tools for their analysis. *Brief Bioinform* 2000;**1**(3):236-49.
9. Kormann MSD, Hasenpusch G, Aneja MK, Nica G, Flemmer AW, Herber-Jonat S, et al. Expression of therapeutic proteins after delivery of chemically modified mRNA in mice. *Nat Biotech* 2011;**29**(2):154-7.
10. Holtkamp S, Kreiter S, Selmi A, Simon P, Koslowski M, Huber C, et al. Modification of antigen-encoding RNA increases stability, translational efficacy, and T-cell stimulatory capacity of dendritic cells. *Blood* 2006;**108**(13):4009-17.
11. Zabner J, Fasbender AJ, Moninger T, Poellinger KA, Welsh MJ. Cellular and molecular barriers to gene transfer by a cationic lipid. *J Biol Chem* 1995;**270**(32):18997-9007.
12. Wilke M, Fortunati E, vandenBroek M, Hoogeveen AT, Scholte BJ. Efficacy of a peptide-based gene delivery system depends on mitotic activity. *Gene Ther* 1996;**3**(12):1133-42.
13. Brunner S, Sauer T, Carotta S, Cotten M, Saltik M, Wagner E. Cell cycle dependence of gene transfer by lipoplex polyplex and recombinant adenovirus. *Gene Ther* 2000;**7**(5):401-7.
14. Mockey M, Goncalves C, Dupuy FP, Lemoine FM, Pichon C, Midoux P. mRNA transfection of dendritic cells: synergistic effect of ARCA mRNA capping with Poly(A) chains in cis and in trans for a high protein expression level. *Biochem Biophys Res Commun* 2006;**340**(4):1062-8.
15. Kuhn AN, Diken M, Kreiter S, Vallazza B, Türeci Ö, Sahin U. Determinants of intracellular RNA pharmacokinetics: implications for RNA-based immunotherapeutics. *RNA Biol* 2011;**8**(1):35-43.
16. De Haes W, Van Mol G, Merlin C, De Smedt SC, Vanham G, Rejman J. Internalization of mRNA lipoplexes by dendritic cells. *Mol Pharm* 2012;**9**(10):2942-9.
17. Arthur JF, Butterfield LH, Roth MD, Bui LA, Kiertscher SM, Lau R, et al. A comparison of gene transfer methods in human dendritic cells. *Cancer Gene Ther* 1997;**4**(1):17-25.
18. Strobel I, Berchtold S, Gotze A, Schulze U, Schuler G, Steinkasserer A. Human dendritic cells transfected with either RNA or DNA encoding influenza matrix protein M1 differ in their ability to stimulate cytotoxic T lymphocytes. *Gene Ther* 2000;**7**(23):2028-35.
19. Seelig G, Soloveichik D, Zhang DY, Winfree E. Enzyme-free nucleic acid logic circuits. *Science* 2006;**314**(5805):1585-8.
20. Rinaudo K, Bleris L, Maddamsetti R, Subramanian S, Weiss R, Benenson Y. A universal RNAi-based logic evaluator that operates in mammalian cells. *Nat Biotechnol* 2007;**25**(7):795-801.
21. Win MN, Smolke CD. Higher-order cellular information processing with synthetic RNA devices. *Science* 2008;**322**(5900):456-60.
22. Leisner M, Bleris L, Lohmueller J, Xie Z, Benenson Y. Rationally designed logic integration of regulatory signals in mammalian cells. *Nat Nanotechnol* 2010;**5**(9):666-70.
23. Xie Z, Wroblewska L, Prochazka L, Weiss R, Benenson Y. Multi-input RNAi-based logic circuit for identification of specific cancer cells. *Science* 2011;**333**(6047):1307-11.
24. Carothers JM, Goler JA, Juminaga D, Keasling JD. Model-driven engineering of RNA devices to quantitatively program gene expression. *Science* 2011;**334**(6063):1716-9.
25. Malone RW, Felgner PL, Verma IM. Cationic liposome-mediated RNA transfection. *Proc Natl Acad Sci U S A* 1989;**86**(16):6077-81.
26. Bettinger T, Carlisle RC, Read ML, Ogris M, Seymour LW. Peptide-mediated RNA delivery: a novel approach for enhanced transfection of primary and post-mitotic cells. *Nucleic Acids Res* 2001;**29**(18):3882-91.
27. Rejman J, Tavernier G, Bavarsad N, Demeester J, De Smedt SC. mRNA transfection of cervical carcinoma and mesenchymal stem cells mediated by cationic carriers. *J Control Release* 2010;**147**(3):385-91.
28. Debus H, Baumhof P, Probst J, Kissel T. Delivery of messenger RNA using poly(ethylene imine)-poly(ethylene glycol)-copolymer blends for polyplex formation: biophysical characterization and in vitro transfection properties. *J Control Release* 2010;**148**(3):334-43.
29. Lin AJ, Slack NL, Ahmad A, George CX, Samuel CE, Safinya CR. Three-dimensional imaging of lipid gene-carriers: membrane charge density controls universal transfection behavior in lamellar cationic liposome-DNA complexes. *Biophys J* 2003;**84**(5):3307-16.
30. Chan C-L, Majzoub RN, Shirazi RS, Ewert KK, Chen Y-J, Liang KS, et al. Endosomal escape and transfection efficiency of PEGylated cationic liposome–DNA complexes prepared with an acid-labile PEG-lipid. *Biomaterials* 2012;**33**(19):4928-35.
31. Safinya CR, Ewert K, Ahmad A, Evans HM, Raviv U, Needleman DJ, et al. Cationic liposome–DNA complexes: from liquid crystal science to gene delivery applications. *Philos Transact R Soc A, Math Phys Eng Sci* 2006;**364**(1847):2573-96.
32. Varga CM, Hong K, Lauffenburger DA. Quantitative analysis of synthetic gene delivery vector design properties. *Mol Ther* 2001;**4**(5):438-46.
33. Dinh AT, Pangarkar C, Theofanous T, Mitragotri S. Understanding intracellular transport processes pertinent to synthetic gene delivery via stochastic simulations and sensitivity analyses. *Biophys J* 2007;**92**(3):831-46.
34. Kamiya H, Akita H, Harashima H. Pharmacokinetic and pharmacodynamic considerations in gene therapy. *Drug Discov Today* 2003;**8**(21): 990-6.
35. Zou S, Scarfo K, Nantz MH, Hecker JG. Lipid-mediated delivery of RNA is more efficient than delivery of DNA in non-dividing cells. *Int J Pharm* 2010;**389**(1–2):232-43.
36. Guo P. The emerging field of RNA nanotechnology. *Nat Nanotechnol* 2010;**5**(12):833-42.
37. Schwake G, Youssef S, Kuhr JT, Gude S, David MP, Mendoza E, et al. Predictive modeling of non-viral gene transfer. *Biotechnol Bioeng* 2009;**105**(4):805-13.
38. Summers HD, Rees P, Holton MD, Rowan Brown M, Chappell SC, Smith PJ, et al. Statistical analysis of nanoparticle dosing in a dynamic cellular system. *Nat Nano* 2011;**6**(3):170-4.
39. Grudzien E, Stepinski J, Jankowska-Anyszka M, Stolarski R, Darzynkiewicz E, Rhoads RE. Novel cap analogs for in vitro synthesis of mRNAs with high translational efficiency. RNA (New York, NY. 2004;10(9):1479–87.
40. Youssef S, Gude S, Rädler JO. Automated tracking in live-cell time-lapse movies. *Integr Biol* 2011;**3**(11):1095-101.
41. de Bruin K, Ruthardt N, von Gersdorff K, Bausinger R, Wagner E, Ogris M, et al. Cellular dynamics of EGF receptor-targeted synthetic viruses. *Mol Ther* 2007;**15**(7):1297-305.

42. Sacchetti A, El Sewedy T, Nasr AF, Alberti S. Efficient GFP mutations profoundly affect mRNA transcription and translation rates. *FEBS Lett* 2001;**492**(1–2):151-5.
43. Blake WJ, Kaern M, Cantor CR, Collins JJ. Noise in eukaryotic gene expression. *Nature* 2003;**422**(6932):633-7.
44. Li X, Zhao X, Fang Y, Jiang X, Duong T, Fan C, et al. Generation of destabilized green fluorescent protein as a transcription reporter. *J Biol Chem* 1998;**273**(52):34970-5.
45. Rädler JO, Koltover I, Salditt T, Safinya CR. Structure of DNA-cationic liposome complexes: DNA intercalation in multilamellar membranes in distinct interhelical packing regimes. *Science* 1997;**275**(5301):810-4.
46. Rädler JO, Koltover I, Jamieson A, Salditt T, Safinya CR. Structure and interfacial aspects of self-assembled cationic lipid−DNA gene carrier complexes§. *Langmuir* 1998;**14**(15):4272-83.
47. Ross PC, Hui SW. Lipoplex size is a major determinant of in vitro lipofection efficiency. *Gene Ther* 1999;**6**(4):651-9.

4 Single-cell array mRNA expression time course analysis for optimization of mRNA stability

Starting from the quantitative study of gene transfer and the mathematical model for pDNA transfection, the characteristic differences of mRNA- and pDNA-mediated protein expression were investigated in P1. There, we found that mRNA transfection leads to fast and homogeneous response and that it is predictive in terms of modeling. In the following study presented in manuscript M1, the dynamics of protein expression induced by different mRNA constructs are assessed and compared.

In gene therapy, the use of mRNA for therapeutic applications has become increasingly popular in recent years. Protein expression induced by mRNA has the advantage that mRNA can not be included into the genome of the target cell because unlike pDNA, mRNA is processed outside the nucleus during translation. This also leads to fast cellular response to mRNA transfection as shown in P1, which is favourable for therapeutic applications and makes mRNA one of the upcoming approaches in regenerative medicine and a potential therapeutic agent for a variety of medical indications [46, 48–50]. Mandal and Rossi recently claimed that "modified mRNA seems poised to emerge as a novel therapeutic paradigm for application in regenerative medicine and beyond" [66]. An ideal therapeutic mRNA construct induces prolonged protein expression above a certain threshold level, so that intervals between repeated administration of the mRNA therapeutic are as long as possible. Novel mRNA constructs show low immunogenic potential due to chemical modification of the nucleotides [49]. In addition to that, a prolongation of their life times can be achieved by inserting untranslated regions into the mRNA sequence [58]. Also, most therapeutic recombinant mRNA constructs have a poly-A tail consisting of 120 nucleotides at their 3' end, which is known to protect the mRNA from degradation [58] (see figure 2.1).

For the study presented in M1, the effect that different combinations of inserted UTRs can cause with respect to the stability and the translational efficiency of the corresponding mRNA construct was systematically investigated using single-cell arrays for mRNA transfection. For each single-cell protein expression time course, characteristic parameters such as mRNA and protein degradation rate and the expression rate $k_{TL} \cdot m_0$ were fitted as described in P1 [54].

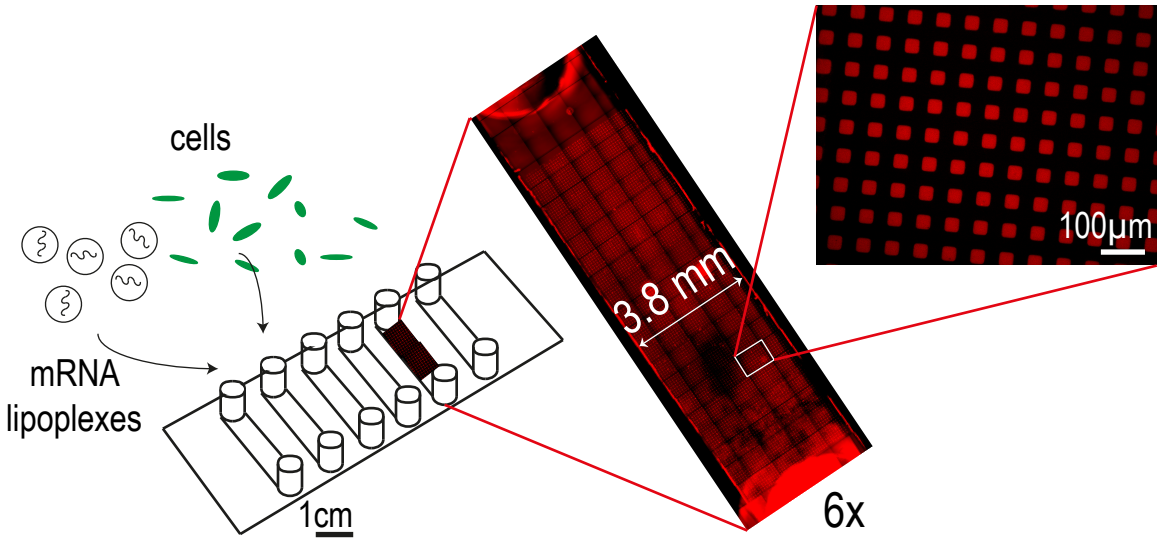

Figure 4.1: *Microstructured multi-channel slide for parallel mRNA transfection on single-cell arrays. Cells were grown on micropatterned single-cell arrays and transfected with mRNA. Enlargements show one out of six micropatterned channels filled with cells and a zoom-in on the protein pattern. To visualise the microstructures, fluorescently labeled fibronectin was used here. Protein expression of d2EGFP after mRNA transfection was monitored using time-lapse fluorescence microscopy. Subsequent image analysis was automated to great extent to generate sufficient statistics at high-throughput. This way, we were able to measure mRNA life time distributions of six different constructs in parallel. All techniques used in these experiments are described in chapter 2.*

Conversion of degradation-rates into life times according to

$$half\text{-}life_{construct_x} = \frac{ln2}{degradation\,rate_{construct_x}} \tag{4.1}$$

yielded single-cell distributions of mRNA and protein life times. This indirect approach circumvents the problem of overestimated mRNA life times that have been observed for measurement techniques such as radioactive labeling, which are based on the total content of transfected mRNA inside a cell, regardless of its actual subcellular location [154].

We choose naturally occuring 5'- and 3'-UTRs of a human gene (namely the *cytochrome b-245 alpha polypeptide, CYBA*) whose mRNA transcripts are known for their long life time. Different combinations of these UTRs were cloned into mRNA constructs encoding for d2EGFP, which was used as a fluorescent reporter protein. For long-term fluorescence time-lapse microscopy, single-cell arrays were prepared as described before [90]. We used multi-channel slides for transfection of up to six different mRNA constructs in parallel as schematically shown in figure 4.1.

Image processing was done according to chapter 2.7.2. This approach enables fast, systematic, and quantitative assessment of mRNA life times at the single-cell level. This way, distinct and characteristic features of the life time distribution of the different constructs can be uncovered. To further corroborate the findings for the analysis of single-cell time courses, additional snap-shot measurements of mRNA-induced protein expression levels were conducted in more cell-types.

The data presented in M1 suggest life time prolonging effects for any UTR-insertion, with the highest values for the 3' - UTR or the 5'+3' - UTR stabilisation. We find this trend in two different cell types. Also, surprisingly low stabilising effects for constructs with a double-insertion of 3' UTRs suggest the impact that the mRNA's secondary structure has on its stability.

The experimental data and the resulting conclusions are presented in the following manuscript M1. This study was a done in collaboration with the group of Prof. Plank at the Technische Universität München. More specifically, all mRNA constructs were cloned and produced by Mehrije Ferizi, who also conducted flow cytometric measurements. Time-lapse fluorescence microscopy and analysis of the corresponding data were done by me.

4.1 Associated manuscript M1

Stability analysis of chemically modified mRNA using micropattern-based single-cell arrays

By

Mehrije Ferizi*, Carolin Leonhardt*, Christian Meggle, Manish K. Aneja, Carsten Rudolph, Christian Plank, Joachim O. Rädler

manuscript in preparation for submission

Stability Analysis of chemically modified mRNA using micropattern-based Single-Cell Arrays

M. Ferizi[1*], C. Leonhardt[2*], C. Meggle[2], M.K. Aneja[3],
C. Rudolph[3], C. Plank[1, 3] J. O. Rädler[2]

[1] Institute of Experimental Oncology and Therapy Research, Technical University of Munich, Munich, Germany

[2] Faculty of Physics and Center for NanoScience, Ludwig Maximilian University, Munich, Germany

[3] Ethris GmbH, Martinsried, Munich

Abstract

The measurement of mRNA turnover in living cells plays an important role in the search for stable mRNA constructs for RNA-based therapies. Here we show that automated time-lapse microscopy combined with micropatterned arrays allows for efficient high-throughput monitoring of fluorescent reporter protein expression at the single-cell level. The fluorescence time courses after mRNA transfection yield the distribution of individual mRNA degradation rates within a population. We compare mRNA constructs with combinations of 5' and 3' UTR sequences and find a systematic broadening and shift towards longer half-lives for UTR stabilized mRNA. Most prominently, the 3' UTR stabilized sequence prolongs life time by a factor of nearly 4 in A549 cells. At the same time the life time distribution of the destabilized EGFP reporter was found to be constant and narrowly distributed. Using mathematical modeling, we show that mRNA life time predicts the area under the curve (AUC) which is relevant for mRNA therapeutics. Our approach paves the way for quantitative assessment of hitherto unexplored mRNA life time heterogeneity, possibly predicated on multiple mRNA secondary structures and its dependence on UTR sequences.

Introduction

In recent years, messenger RNA (mRNA) has become increasingly relevant as a new drug entity. As opposed to DNA-based gene therapeutics, mRNA does not need to be transported into the nucleus but is directly translated into protein in the cytoplasm (1, 2). This makes mRNA safer in avoiding potential insertional mutagenesis, an unlikely but existent risk of DNA gene medicines. As a consequence, mRNA therapeutics are emerging as promising alternatives for gene and protein replacement therapies in a broad variety of medical indications (1-4). However, the strong immunogenicity as well as the limited stability of conventional mRNA has to be overcome to further establish its clinical applicability. In particular, mRNA stability is an essential parameter for envisaged medical applications because it determines, for example, dosing and the dosing intervals of mRNA drugs.

Several strategies have proven successful both at increasing the stability and reducing the immunogenic response triggered by mRNA administered to cells or organisms. Amongst these is the inclusion of chemically modified nucleotides (5). Kormann et al. have shown that the replacement of only 25% of uridine and cytidine residues by 2-thiouridine and 5-methyl-cytidine suffices to increase mRNA stability as well as to reduce the activation of innate immunity triggered by externally administered mRNA in vitro (3).

Also, untranslated regions (UTRs) in mRNAs have been reported to play a pivotal role in regulating both mRNA stability and mRNA translation. UTRs are known to influence translational initiation, elongation, and termination, as well as mRNA stabilization and intracellular localization through their interaction with RNA binding proteins (6, 7). Depending on the specific motives within the UTR, it can either enhance or decrease mRNA turnover (8-11). Recently, data on mRNA half-lives and the corresponding UTR sequences have been published (12).

One mRNA with a particularly long half-life is the one transcribed from the human *cytochrome b-245 alpha polypeptide* (CYBA) gene. The CYBA gene comprises specific 5' and 3' UTRs. In general, 5' UTR motives such as upstream open reading frames (uORFs) or internal ribosomal entry sites (IRES) are known to be involved in gene regulation, particularly in translational initiation (13). The 3' UTRs can comprise even more regulatory functions than the 5'UTRs, some of them even hindering mRNA translation (14). While no regulatory motives are known for the CYBA 5' UTR unit, the CYBA 3' UTR contains two of them. Firstly, the polyadenylation signal (PAS), which interacts with the cytoplasmic polyadenylation element binding protein (CPEB), as well as with the cleavage and polyadenylation signaling factor (CPSF) (11). CPEB is responsible for the prolongation of the poly-A tail in the cytoplasm, whereas CPSF primes the pre-mRNA through cleavage at a specific site for the

upcoming addition of poly-A (11, 14). As a second regulatory motif, the CYBA 3' UTR contains the insulin 3' UTR stability element (INS_SCE). The INS_SCE sequence has been shown to bind to the polypyrimidine tract binding protein (PTB) under reducing conditions, increasing the mRNA half-life of insulin (15). UTRs of CYBA are shown in Supplementary Table S2. Another important feature influencing mRNA translation efficiency is the poly-A tail, which is located on the 3' end. It has been shown that a prolongation of the poly-A tail to 120 nucleotides has beneficial effects on protein expression, assumingly because of the protective effect of longer poly-A tails against mRNA degradation (16). In contrast to long poly-A tails, mRNAs with poly-A tails shorter than 50 nucleotides are claimed not to be translated at all (11, 17). Hence, in mRNA therapy, recombinant mRNA constructs are likely to be furnished with a poly-A tail of 120 nucleotides. Degradation of most mRNA transcripts in eukaryotic cells begins with 3' to 5' exonucleolytic deadenylation, resulting in removal of most of the poly A-tail. Subsequently, two major pathways that are responsible for the degradation of the rest of the mRNA body are known to come into play. On the one hand, the 5' end is decapped by the Dcp1/Dcp2 complex, followed by 5'-3' exonucleolytic degradation that is catalyzed by Xrn1p. On the other hand, the exosome enables 3'-5' exoribonucleolytic degradation with the 5' cap being retained (18). Moreover, it is known that the 5' cap interaction with the 3' poly-A tail results in circular forms of the mRNA. It is assumed that the circular shape of the mRNA increases the initiation rate of ribosomes after translating the first stop codon and also protects mRNA against degradation (19).

We were interested in whether the reported long half-life of natural CYBA mRNA can be conferred to a foreign mRNA by virtue of flanking its coding sequence with combinations of CYBA 5'- and 3'-UTRs. So far, mRNA stability and turnover has been measured by a variety of approaches including metabolic labelling, mass spectrometry, quantitative real time reverse transcription polymerase chain reaction, microarrays, or fluorescence labelling techniques for imaging mRNA inside cells (16, 20-25). Most of these are ensemble measurements.

Recently, we have shown that single-cell analysis of mRNA transfection time-lapse movies is capable of assessing individual expression time courses yielding the mRNA decay rates at the single-cell level (26). Furthermore, we have reported on the use of regular micropatterns to position cells on a regular grid of adhesion sites (27). Hence, we were interested in whether this technology is suited to rapidly analyse the functions of CYBA UTR combinations on a foreign mRNA. To address this question, we have chosen the coding sequence of destabilized enhanced green fluorescence protein (d2EGFP) to artificially shorten the life cycle of the reporter protein inside the cell (28). The combinations included insertion of the

respective CYBA UTRs at 5' or 3' ends, respectively, at both 5'- and 3' ends, at the 5' end combined with two repeats of the 3' UTR at the 3' end, or two repeats of 3' UTR without 5' UTR, and all of these compared to a control construct without UTRs. Protein and mRNA life times and the expression rate from each of the compared transcripts were assessed. The dynamics of gene expression after mRNA transfection were compared to a population based method (FACS and fluorescence microscopy). We observed an improvement of total protein expression over a period of three days for all UTR combinations compared to the control.

Results

Fluorescence microscopy and analysis via flow cytometry (FC)

To evaluate the effect of different UTR combinations on transgene expression kinetics, two different cells lines were transfected using Lipofectamine^{TM}2000 with different d2EGFP mRNA constructs containing a 5' UTR alone, a 3' UTR, 5'+3' UTR, two copies of 3'UTR and 5'+2x3' UTR. A schematic representation of the building blocks of all constructs can be seen in Figure 1A. At different time points through three days post-transfection, d2EGFP expression was quantified using FC. An exemplary dot plot for t=24h, illustrating d2EGFP expression levels of live A549 cells, is shown in Figure 1C (see Figure S4 B for corresponding Huh7 data). In addition, we imaged the cells using fluorescence microscopy (see Figure 1B and D and Figure S4 A and C). Comparable transfection efficiencies for all mRNA constructs were confirmed 24 hours post transfection (Figure 1B and S5 A). Thereby, we can rule out differential transfer efficiencies to be a causal factor for the observed differences in expression kinetics. Based on fluorescence microscopy images, we detected a drastic reduction of d2EGFP expression for all constructs at 48 h post-transfection (see Figure 1B and1D and S4 A and C). However, higher EGFP expression levels respect to the control were found for all UTR-stabilized mRNAs. More specifically, mRNA constructs containing 3' UTRs seemed to enhance expression more than constructs without 3' UTRs. We observed this for A549 and Huh7 cells (see Figure 1 and S 4, respectively). At time points later than 48h, this effect was pronounced even more (data not shown). In Figure 2 A and B, the time courses of the mean fluorescence intensities (MFI) as determined by FC are shown for all constructs in both cell types. Also here, all UTR-containing mRNA constructs showed higher MFI values than the control construct in both cell lines at all points in time. Taken together, our fluorescence microscopy and FC data suggest that mRNA molecules furnished with CYBA UTRs show persistent d2EGFP expression for more than 24 hours.

Single-cell expression arrays

We fabricated microstructured, cell-adhesive substrates as shown in Figure 3A and B as a platform for single-cell time-lapse microscopy. The rectangular squares are functionalized with the extracellular matrix protein fibronectin, while the surrounding dark area is passivated with cell repellent PEG-PLL. Cells were seeded at an appropriately dilute cell density such that after about three hours, cells adhered to the rectangular squares. This cellular self-organization process has been studied in detail before (27). The size of the squares was 30µm for optimal filling with single cells. The distance between the squares was just big enough (60µm) to minimize bridging effects of cells adhering to more than one square at the same time. Time-lapse fluorescence microscopy and automated image analysis of the fluorescence signal per square yields hundreds of individual time courses. A typical set of background corrected raw data is shown in Figure 3D. The red lines represent exemplary fits to the mathematical expression for mRNA translation (see also Materials and Methods section). Data were analyzed as described recently (26) by fitting each time-course with the analytical solution for mRNA-induced protein expression,

$$G_{mRNA}(t) = \frac{K}{\delta - \beta} \cdot \left(1 - e^{-(\delta-\beta)(t-t_0)}\right) \cdot e^{-\beta(t-t_0)} \quad \text{(Equation 1)}$$

using IgorPro software. Here, G denotes the amount of protein, K is the expression rate, δ is the mRNA degradation rate, and β is the degradation rate of the reporter protein d2EGFP. The expression rate K is the product of the initial amount of mRNA molecules inside the cell (m_0) and the translation rate k_{TL}. The time-course that is described by Equation 1 will be discussed in detail in below section “mastercurves of protein expression”.

In vitro transfection on cell arrays

In a typical experiment, cells were allowed to adhere to the micropatterns for three hours before transfection. Each of the six microchannels was filled with a different lipoplex solution, containing one of the constructs of interest. In initial experiments, we compared two different, commercially available transfection reagents (namely Lipofectamine™ 2000 and DOGTOR). We found higher transfection efficiencies for Lipofectamine™ 2000 than for DOGTOR (see Figure S1). Because we additionally obtained high cell viability rates of above 80% with Lipofectamine™2000 (data not shown), all further transfection experiments were conducted using Lipofectamine™2000. As mRNA-mediated protein expression starts shortly after transfection, incubation time was kept to a minimum. Accordingly, the ratio between mRNA dosage and incubation time was adjusted to achieve high transfection efficiencies (see also Supplementary, Figure S1) and negligible toxic effects caused by over-expression of the reporter protein. At an mRNA dose of 5pg/cell, an incubation time of one hour was found to be optimal.

Expression rates

All results for the two cell types are based on four independent measurements under the same experimental conditions. Time-lapse data of about thousand A549 cells and thousand Huh7 cells have been analyzed. The distributions of the obtained expression rates K are shown in Figure 4A and the corresponding mean values can be seen in Figure 4D. Both the mean expression rates and the shape of their distributions were found to be rather similar for the different constructs.

mRNA half-lives

We converted the fitted mRNA-degradation rates δ into mRNA half-lives according to

$$\tau = \frac{ln2}{\delta}. \qquad \text{(Equation 2)}$$

Figure 4B shows the half-life distributions of differently stabilized mRNA constructs in A549 and Huh7 cells, respectively. Here, it becomes evident that for stabilized constructs, both mean half-life and broadness of the underlying distribution increase as compared to the reference construct.

An overview of all determined half-lives is given in Figure 4D. Both for A549 and for Huh7 cells, we found the longest half-life for mRNA stabilized either by a 3'UTR alone or both 5'and 3'UTR. The life time prolonging effect is more pronounced in A549 cells. Expectedly, the shortest half-life was obtained for the control construct that does not contain a stabilizing UTR. It amounts to 5.8 hours for A549 cells and to 7.8 hours for Huh7 cells.

Protein half-lives

The distributions of protein (d2EGFP) degradation life times are presented in Figure 4C. As expected the half-lives of the expressed protein do not vary for the different mRNA constructs. The determined mean life times range from 4.2 to 4.9 hours for A549 cells and from 5.6 to 8.5 hours for Huh7 cells as shown in Figure 4D. The coefficients of variation are about 0.29 (A549) and 0.45 (Huh7) and hence is significantly smaller than the coefficient of variation of up to 0.6 that we found for the distribution on mRNA life-times. As a control, we also measured the half-lives in an alternative approach, where translation was inhibited by addition of cycloheximide at a given time point, t_0, after transfection (see Figure S3 in the Supplement). In this case, protein expression is induced for a while and then stopped. The exponential decay in fluorescence after inhibition yields protein life times. These half-lives were found to be smaller by a factor of about two, compared to the above experiments without inhibition. In both experiments, however, the relative ratios of the protein life times in Huh7 cells as compared to those in A549 cells is the same.

Mastercurves of protein expression

The features of mRNA induced protein expression become evident in the so-called mastercurve of protein expression as depicted in Figure 5A (A549) and B (Huh7). The mastercurve is the population average of the onset-time corrected single cell traces, i.e. all onset-times were shifted to time point zero. Fluorescence intensities were converted into actual numbers of d2EGFP as described before in reference (26).

The superior properties of the 3' and the 5'+3'-stabilized mRNA constructs are illustrated in the mastercurve plot. These constructs showed the shallowest decrease in protein expression with time and hence the longest half-lives in addition with higher protein expression values as compared to the other constructs.

Area under the curve (AUC)

An optimal therapeutic mRNA construct should desirably have both long mRNA half-life and high translational efficiency. The integral of the amount of protein over time, i.e. the area under the curve (AUC), is a means to simultaneously quantify the translational efficiency and the stability of an mRNA construct. It can be interpreted as the cumulative time-dose of the protein that is encoded on the mRNA and hence describes the efficacy of a chosen mRNA construct. Given the biochemical model, the AUC can be calculated explicitly as follows (see also Supplementary):

$$AUC = \frac{K}{\delta \cdot \beta} \qquad \text{(Equation 3)}$$

An illustrative explanation for the calculated AUC can be seen in Figure 6A. If there was no protein degradation (β=0), the amount of protein inside a cell would run into a steady state level as a consequence of a balanced flux of mRNA translation and mRNA degradation. In this case the expression dynamics follows $\frac{K}{\delta}(1 - e^{-\delta t})$. The same would be true in an analogous manner for the case where δ was equal to zero. The superposition of this with the permanent, exponential decay of the d2EGFP protein (following $e^{-\beta t}$) results in the characteristic shape of the AUC as shown in Figure 6A. Figures 6B and C show the overall mean relative AUCs as well as the "per-experiment" relative AUCs with respect to the mean AUC of the control, the latter being the AUC of protein expression after transfection with the control construct. In both cell types, we find the highest relative AUCs for the 3'UTR- and the 5'+3'UTR-stabilized construct. This is consistent with the observed long half-lives for these constructs, because they contribute to the AUC as seen in equation 3. The detailed, single-cell AUC distributions can be found in Figure S2 of the Supplementary.

Life time-prolongation factor

The life time-prolongation factors for A549 and Huh7 cells are shown in Figure 5D and E, respectively. As expected, all stabilized constructs yield life time-prolongation factors higher than one, meaning that the insertion of UTRs at either end causes mRNA stabilization. However, the 3'UTR mRNA construct shows longer mRNA life times than the 2x3'UTR construct. Similarly, the 5'+3'UTR construct is more stable than the 5'+ 2x3' construct. These results hold true for both cell types. Interestingly, the stabilizing effects are significantly more pronounced in A549 cells than in Huh7 cells in all cases.

Discussion

Determination of mRNA stability and its expression are two major factors to be considered when it comes to developing new mRNA therapeutics. Here, we used different combinations of UTRs, a 5' UTR, 3'UTR, a 5'+3' UTR, 5'+2x3' UTR, and two copies 3' UTR to improve mRNA in terms of stability and its expression. The AUC of the d2EGFP time course is also evaluated, because the total protein expression is relevant for a sustained therapeutic effect. In order to get detailed time-resolved data and monitor protein expression dynamics at the single-cell level, we used microstructured single-cell arrays for parallel, quantitative measurements of mRNA stability and translational efficiency. The regular arrangement of cells guaranteed reproducible microenvironments and enabled fast and automated image-analysis, which are prerequisites for comparative, high-throughput single-cell studies. The approach allows the determination of distribution functions for (i) protein half-life, (ii) expression rates, and (iii) mRNA half-life.

In both A549 and Huh7 cells, mean protein half-lives of d2EGFP were narrowly distributed and independent of the UTR sequence. The calculated half-life values of 4.5 hours for A549 cells and 7.4 hours for Huh7 cells could be attributed to cell type specific differences between the compared cell lines. Such cell specific differences in d2EGFP half-life have been published previously. A study in NIH3T3 cells using a similar imaging cytometry approach, recorded a half-life of 2.8 h within a measurement window of 12 hours (29). An even shorter half-life of less than two hours has been reported for CHO cells by Li et al.(30). Here, protein degradation was measured by Western blotting and flow cytometry for three hours only.

To validate our findings from single-cell data analysis, we additionally determined d2EGFP life times in direct measurements using cycloheximide (see Supplementary, Figure S3). We found shorter life times as compared to the values observed from single-cell data analysis. This might be due to the fact that in single-cell data analysis, a constant initial number of mRNA molecules was assumed as part of the combined expression rate $K=k_{TL}*m_0$ (see Equation 1). However, regardless of the fact that cells have been washed after one hour incubation time, it is still likely that the number of mRNA molecules is not constant from the

start of observation. As a consequence, mRNA molecules that are available for translation later on, leading to protein expression, might result in longer half-life values obtained from single-cell expression time course fitting. When we compare the mean half life determined for A549 cells with the mean half life determined for Huh7 cells, we find the same ratio of roughly 1.64 for both measurement methods. Also, even a possible systematic over-estimation of mRNA and protein half-lives does not change the qualitative order of the mRNA performance as determined by us.

The expression rate depends on the initial number of mRNA molecules, m_0, as well as on the translation rate K_{TL}. Note that the number of successfully delivered mRNA molecules varies due to the intrinsic stochasticity of the delivery process. The mean number of mRNA molecules, however, is expected to be the same, since the transfection protocol has scrupulously been kept up in all experiments. In contrast, the translational activity (K_{TL}) of the various UTR constructs might vary. Still, the fact that the distributions as well as the mean values of the expression rate K are rather similar for all constructs (see Figure 3A and D) indicates that the translation rate is merely influenced by the inserted UTRs.

The parameter of highest interest is the mRNA half-life. Our results suggest that any insertion of 5' or 3' UTRs into the mRNA sequence increases its stability. All modifications tested in this study led to prolonged mRNA half-lives (see Figures 2 and 3) thereby resulting in prolonged expression as measured by fluorescence microscopic imaging and FACS (see Figure 1). The heterogenic distribution of the half-lives points out the importance of single-cell measurement techniques, because these effects are obscured in ensemble measurements (Figure 2 and Figure 4). Interestingly, a positive effect on protein expression was observed for 5' UTR alone, although till date, no known motif in the CYBA 5' UTR has been discovered. For the first time, we could show that CYBA UTRs at either end suffice to increase both peak and persistent of protein expression in both cell lines. These findings are consistent with publications claiming individual or synergistic behaviour of 5' UTRs and 3' UTRs (14). In contrast to Holtkamp et al. (16), no additional increase in protein expression or mRNA stability could be observed with two sequential copies of the 3'UTR as compared to one single 3' UTR (see Figure 4). Conversely, it even resulted in shorter life times both for 5'+3' versus 5'+2x3' UTR insertion and for 3' versus 2x3' UTR insertion. This might be due to the fact that a different type of cells (namely dendritic cells) was used in the study by Holtkamp et al. (16). Another contributing factor affecting both mRNA stability and its translation efficiency might be the secondary structure of the different mRNAs. Such effects of mRNA secondary structure in regulating gene expression have been reported before (31, 32).

Important structural characteristics together with their minimum free energy for the mRNA constructs used in the current study are summarized in supplementary Table 1. The

persistent protein expression of the 5’ + 3’UTR stabilized construct could be due to binding of the 5’ to the 3’end, which facilitates circularization of the mRNA (19). Because we could not find any stable secondary structures within the 5’ UTR, we assume that this feature enables an early expression onset (33). In contrast, we identified secondary structures within the 3’ UTRs. These might protect the mRNA from the 3’-5’ degradation pathway. Two 3’ UTRs showed even more secondary structures (two hairpins) with the best minimum free energy, indicating more persistent expression and slower mRNA degradation at the cost of slow onset of translation and reduced peak expression. Taken together, these findings could be the explanation for the inferior onset expression of the 2x3’ UTR compared to the 5’UTR and the persistent expression at later time points of mRNA constructs containing 3’ UTRs.

In accordance with protein half-lives, longer half-life values were obtained for mRNAs stabilized with UTRs. This was observed in both cell lines with cell specific differences most likely affecting the absolute values. In A549 cells, mRNA half-lives for the constructs with UTRs ranged from 13.0 h to 23.0 h as compared to 5.8 h for the control. In Huh7 cells, half-lives from 9.9 h to 13.6 h were measured for UTR-containing constructs, as opposed to a half-life of 7.8 h for the control mRNA. The half-life of the 3’UTR-stabilized mRNA in A549 cells is in good agreement with mRNA life times of similarly stabilized mRNAs that were reported previously (16, 26). The fact that stability and decay kinetics of mRNA and protein differ in different cell types is most likely due to differences in the complex networks of interactions between mRNA and proteins which are very likely to be cell-type dependent.

Taken together, our results in both A549 and Huh7 cells, independent of the analysis method (FACS or single-cell analysis), suggest that sustained, high levels of protein expression can be induced by CYBA UTR stabilized mRNA. The choice of UTR combination depends on the need of the experiment of application. Where persistent protein expression with reduced mRNA decay is desired, mRNA stabilized with a 3' UTR alone might serve the purpose. However, the combination of 5'+3' UTR results in additional desirable features of early onset, high peak and cumulative protein expression.

Outlook

Our study demonstrates that single-cell analysis of mRNA-induced protein expression is a means to characterize and improve pharmacokinetic properties of mRNA constructs. Using this approach, we were able to systematically assess the intracellular bioavailability of different mRNA constructs to identify sequences yielding sustained protein expression. In our single-cell model we found most prolonged persistence of protein expression for constructs stabilized either by 3’UTR alone or 5’+3’UTR in two cell types. We hope that our study helps to further refine strategies for the manipulation of mRNA stability for future pharmaceutical applications.

Materials and Methods

Plasmid Vectors

Destabilized Enhanced Green Fluorescent Protein (d2EGFP) was excised from pd2EGFP-N1 (Clonetech) and cloned in pVAXA120 (3) to generate pVAXA120-d2EGFP. Based on previously published data with respect to mRNA stability, preselected 5' and 3' UTR sequences of *cyba* gene were synthesized by Eurofins MWG (Germany) and cloned upstream (5'UTR) and/or downstream (3'UTR or 2x3'UTR) of d2EGFP in pVAXA120-d2EGFP, thereby generating the constructs with respective UTR combinations.

mRNA Production

To generate *in vitro* transcribed mRNA (IVT mRNA), plasmids were linearized downstream of the poly(A) tail by *Not*I digestion and purified by chloroform extraction and ethanol precipitation. Purified linear plasmids were used as template for *in vitro* transcription using RiboMax Large Scale RNA production System-T7 (Promega, Germany). Anti-Reverse Cap Analog (ARCA) was added to the reaction mix to generate 5' capped mRNA. Additionally for the production of SNIM mRNAs, chemically modified nucleotides namely methyl-CTP and thio-UTP (Jena Bioscience, Germany) were added to a final concentration of ATP:CTP:UTP:methyl-CTP:thio-UTP:GTP of 7.57mM:5.68mM:5.68mM:1.89mM:1.89mM:1.21mM. The complete IVT mix was incubated at 37°C for 2 hours following which, DNA template was digested with *DNase*I. RNA was precipitated with ammonium acetate (final concentration 2.5M) and washed with 70% EtOH. The washing step was performed twice. Finally, the RNA pellet was re-suspended in RNAse-free water. All mRNA transcripts were verified on 1% agarose gels. A schematic representation of an exemplary mRNA construct can be seen in Fig. 1A. The exact sequences of the UTRs are given in the Supplementary (table S1).

Flow Cytometry (FC)

The experimental set-up looks like as follows: 20,000 cells in 150µl medium were seeded per well in 96-well plates and transfected 24 hours post-seeding. Cells were transfected at a dose of 5pg mRNA/cell using the commercial transfection reagent Lipofectamine^{TM}2000. Complexes were prepared at a ratio of 2.5µl Lipofectamine^{TM}2000 per 1 µg mRNA. For the formation of lipoplexes, Lipofectamine^{TM}2000 and mRNA were diluted separately in OptiMEM transfection medium in a total volume of 50µl, each. These mixtures were incubated at room temperature for 5 minutes. The mRNA solution was then mixed with the Lipofectamine^{TM}2000 solution, followed by another 20 minutes of incubation at room

temperature. After incubation, 900µl of OptiMEM were added to the lipoplex solution. Finally, 50µl of the complex solution were added to the cells and incubated for 1 hour. For every mRNA construct, biological triplicates were prepared. After incubation, the lipoplex-solution was discarded and fresh 150µl medium was added to each well. d2EGFP expression was measured after 8, 24, 36, 48, 60 and 72 hours using FC. Fluorescence microscopy images were taken at each of these time points. For FC measurements, the cell culture medium was discarded and the cells were washed with 1xDPBS (Gibco Life Technology). Subsequently, 20µl of TrypLE Express (Gibco Life Technology) were added per well and incubated for 5 min at 37°C. The reaction was neutralized by adding 80µl 1xPBS, supplemented with 2% FBS. Cells were mixed by pipetting and were transferred into a 96 well plate appropriate for flow cytometric measurements. Finally, 5µl of Propidium iodide (final concentration 1µg/ml) were added per well and measured with Attune Auto Sampler (Applied Biosystems). Please note that fluorescence images were taken prior to FC analysis with a JULY™ microscope.

Surface Patterning and Sample Preparation

Microstructured surfaces were produced by selective oxygen plasma treatment (Femto Diener, 40 W for 3 min) on a top as substrate (ibidi GmbH) with subsequent passivation. Selectivity was achieved using a polydimethylsiloxane (PDMS) stamp (cast from a master produced by photolithography) as a mask. The parts exposed to plasma were passivated by incubation for 30 min with PLL(20k)–g(3.5)-PEG(2k) at a concentration of 1 mg/ml in aqueous buffer (10 mM HEPES pH 7.4 and 150 mM NaCl). Thereafter, the samples were rinsed with PBS and the PDMS stamps were removed. The foils were then fixed to adhesive six-channel slides (sticky µ-slide VI). Each channel was filled with a solution of 50 µg/ml fibronectin in PBS for one hour to render the remaining sectors cell-adhesive. Probes were thoroughly rinsed with PBS three times. The samples were stored in cell medium at room temperature before cell seeding. For this study, square adhesion sites of 30 µm x 30 µm were used because this size turned out to be reasonable for single-cell adhesion of A549 as well as Huh7 cells. Cells were seeded at a density of 10,000 cells per channel so that roughly one cell could adhere on each cell-adhesive island. To obtain fluorescent micropatterns as shown in Figure 3A, a mixture of 20 µg/ml fibronectin and 30 µg/ml fibrinogen conjugated with Alexa Fluor 488 was used.

Materials

FBS, Leibovitz's L-15 Medium (Gibco), Lipofectamine™2000, and OptiMEM (Gibco) were purchased from Invitrogen, Germany. Sterile PBS was prepared in-house. Ham's F-12K, DMEM, and Trypsin-EDTA were purchased from c.c.pro GmbH, Germany. Channel slides were purchased from ibidi, Germany. Fibronectin was purchased from Yo Proteins, Sweden.

PLL-g-PEG was purchased from SuSoS AG, Switzerland. Alexa Fluor 488 was purchased from Life Technologies, Germany. The plasmid pd2EGFP-N1 was purchased from BD Biosciences Clontech, Germany.

Cell Culture

A human alveolar adenocarcinoma cell line (A549, ATCC CCL-185) was grown in Ham's F12K medium supplemented with 10% FBS. A human hepatoma epithelial cell line (Huh7, JCRB0403, JCRB Cell Bank, Japan) was cultured in DMEM medium, supplemented with 10% fetal bovine serum. All cell lines were grown in a humidified atmosphere at 5% CO_2 level.

***In vitro* Transfection**

Three hours prior to transfection, 10.000 cells per channel were seeded in a 6-channel slide. Cells were transfected at a dose of 5pg mRNA/cell using the commercial transfection reagent Lipofectamine^{TM}2000at a ratio of 2.5µl Lipofectamine^{TM}2000per 1 µg mRNA. The complex formation was prepared as follows: Lipofectamine^{TM}2000 and mRNA were separately diluted in OptiMEM transfection medium to add up to a total volume of 45µl, each. These mixtures were incubated at room temperature for 5 minutes. The Lipofectamine^{TM}2000 solution was then mixed with the mRNA solution, followed by another 20 minutes of incubation at room temperature. Please note that the microchannels were never empty during all subsequent rinsing steps: Immediately before transfection, the cells were washed with PBS. Finally, the lipoplex solutions containing different mRNAs constructs were filled into the six channels. All five different mRNA constructs plus the reference construct could thus be measured under the same experimental conditions. The cells were incubated in a total transfection volume of 90µl at 37°C (5% CO2 level) for one hour. The transfection medium was thereafter removed and the cells were washed with PBS. Subsequently, the cells were re-incubated with Leibovitz's L-15 Medium containing 10% FBS before monitoring of d2EGFP expression.

Data Acquisition and Quantitative Image Analysis

Live-cell imaging was performed on a motorized inverted microscope (Nikon, Eclipse Ti-E) equipped with an objective lens (CFI PlanFluor DL-10×, Phase1, N.A. 0.30; Nikon) and with a temperature-controlled mounting frame for the microscope stage. To acquire cell images, we used a cooled CCD camera (CLARA-E, Andor). A mercury light source (C-HGFIE Intensilight, Nikon) was used for illumination and a filter cube with the filter set 41024 (Chroma Technology Corp., BP450-490, FT510, LP510-565) was used for d2EGFP detection. An illumination shutter control was used to prevent bleaching. Images were taken at 10 fold magnification with a constant exposure time of 600 ms at 10 minute-intervals for at least 25 hours post-transfection. Fluorescence images were consolidated into single-image sequence files.

Quantitative analysis of characteristic parameters of single-cell expression kinetics allows the comparison of various vector performances in terms of expression efficiency and stability.

Image analysis consisted of several steps and was done using in-house-developed software based on ImageJ. First, a rectangular grid was overlaid with the original time-lapse movie and adjusted to the size and orientation of the underlying cell-pattern. Next, the software automatically detected d2EGFP-expressing cells by reading out the fluorescence intensities of all squares. Unoccupied squares were used for background correction. The software calculates the cells' fluorescence over the entire sequence and connects corresponding intensities to time courses of the fluorescence per cell. Finally, single-cell fluorescence intensities per square were extracted.

Data were then analyzed as described recently by fitting each time-course with the analytical solution for mRNA-induced protein expression,

$$G_{mRNA}(t) = \frac{K}{\delta - \beta} \cdot \left(1 - e^{-(\delta-\beta)(t-t_0)}\right) \cdot e^{-\beta(t-t_0)}, \qquad \text{Equation 1}$$

using *IgorPro* software. Here, G denotes the amount of protein, K is the expression rate, δ is the mRNA degradation rate, and β is the degradation rate of the reporter protein d2EGFP. A schematic representation of the simplistic model assumed for mRNA-induced protein expression can be seen in Figure 3C.

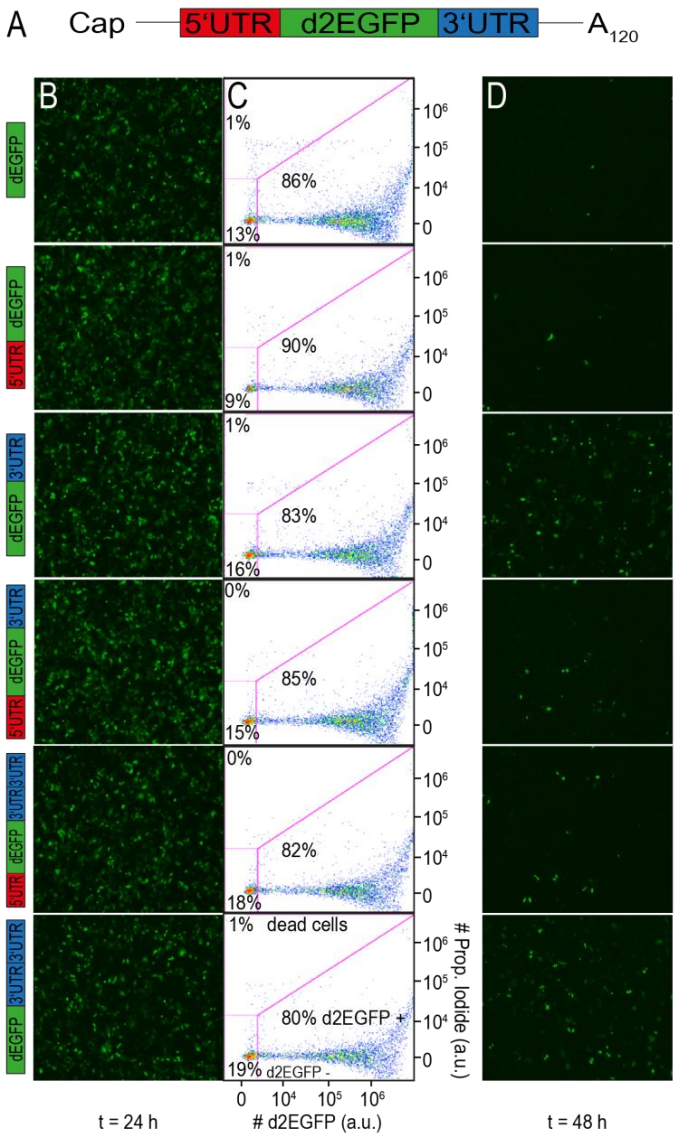

Figure 1 **Fluorescence microscopy and flow cytometry data of A549 cells**

A) Schematic illustration of therapeutic mRNA, consisting of a 5' CAP, a 5' UTR, an encoding region, a 3' UTR and a poly-A tail.

B) Fluorescence microscopy pictures taken with 4-fold magnification (JULY^{TM}) at 24 hours post-transfection. All constructs showed improved protein expression levels as compared to the control.

C) The percentage of d2EGFP positive cells as determined by FC is similar for all constructs. Propidium iodide was used to detect dead cells. The applied gates ensured exclusion of dead cells and untransfected cells.

D) At 48 hours post transfection, sustained protein expression was higher for the stabilized constructs as compared to the control.

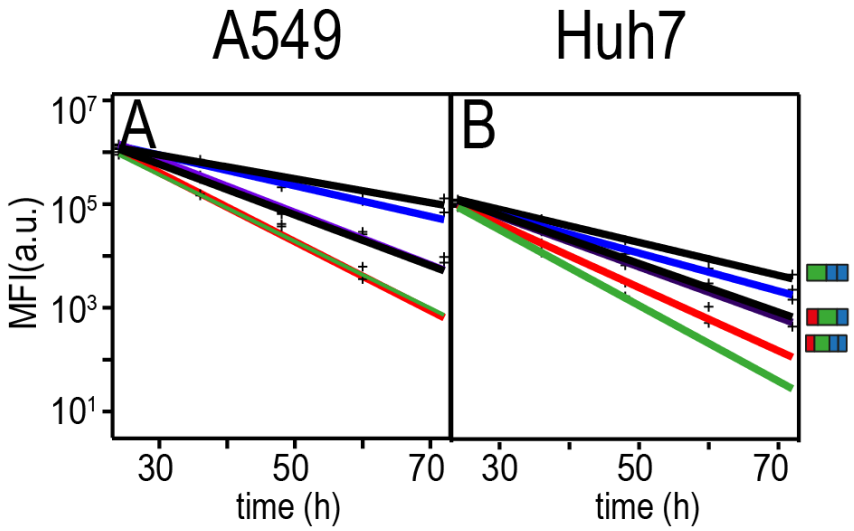

Figure 2 **Time courses of protein expression as determined by FC**

Mean fluorescence intensities are plotted versus time in a semi-logarithmic plot. With increasing time post transfection, the elevated protein expression levels of the stabilized constructs become more and more evident. The green, red and blue lines correspond to the control/5'UTR/3'UTR constructs, respectively. The black lines correspond to the constructs that are shown on the right hand side in the order of appearance.

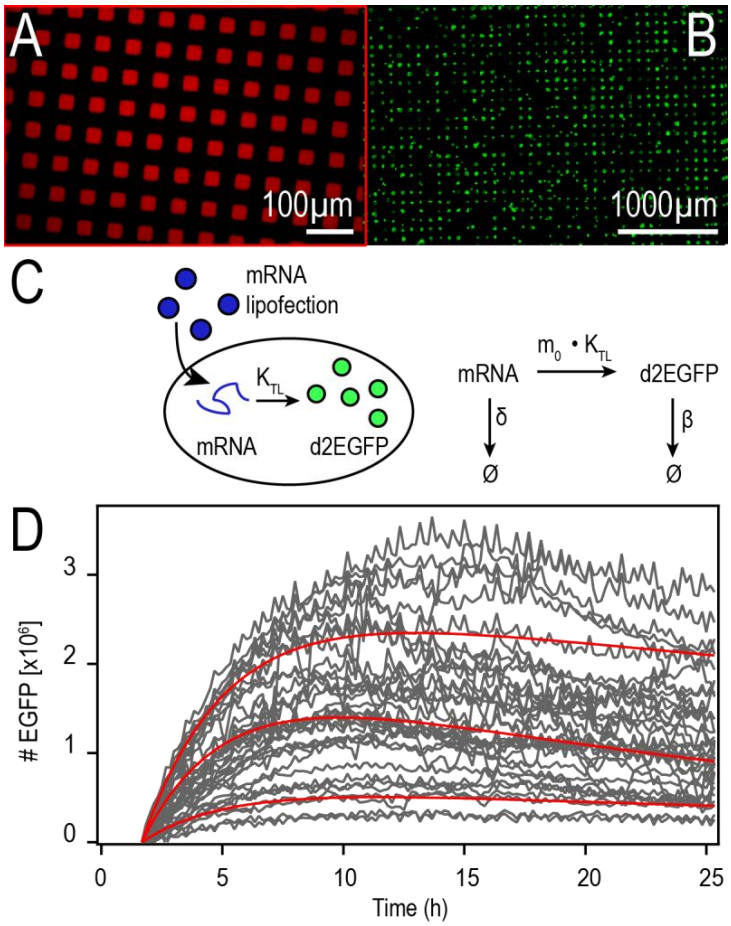

Figure 3 **Microstructured multi-channel slides** for parallel single-cell assays to test differently stabilized mRNA constructs

A) Cell-adhesive, microstructured protein patterns with cell-repellent PEG areas in between allow ordered cell arrangement. Fluorescently labeled fibronectin was used to visualize the micropattern.

B) Fluorescent A549 cells adhering to fibronectin patterns inside a microchannel (three hours after seeding). C) Schematic drawing of mRNA lipofection (on the left) and reaction scheme underlying our analytical solution (on the right).

D) Exemplary time courses of mRNA-mediated d2EGFP expression in A549 cells. Red lines are representative fits to the theoretical translation model.

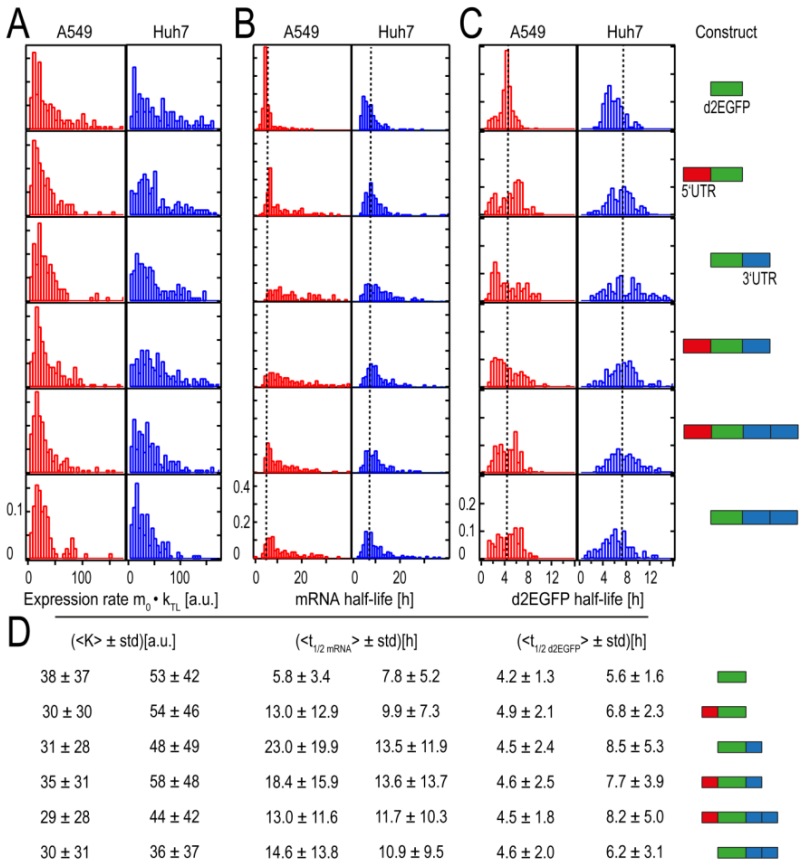

Figure 4 **Distributions of expression rates K, mRNA-Life times, and d2EGFP-life times and corresponding mean values with schematic representations of the constructs**

A) Distributions of expression rate K, which is the product of the initial number of mRNA molecules and the translation rates. The fact that the distributions are similarly shaped indicates that the transfection kinetics and the translation rates are very similar.

B) The distributions of the mRNA half-lives show great variations in their broadness. As a guide to the eye, dotted lines indicate the mean half-life of the control.

C) Distributions of d2EGFP half-lives. As expected, the distributions of the different constructs are similarly shaped and show comparable mean values. As a guide to the eye, the overall mean half-life of d2EGFP based on all measured half-lives is shown as a dotted line.

D) Mean values and the corresponding standard deviations (std) of the fitted rates. Although the control construct yields high mean K values in both cell types, the short mRNA half-life of this construct leads to small AUC values as compared to the stabilized constructs. This can be seen in Figure 6.

Schematic representations of the constructs can be seen on the right hand side. All constructs have the same 5'cap and a poly(A) tail. Data from 895 single A549 and 1355 Huh 7 cells were analysed.

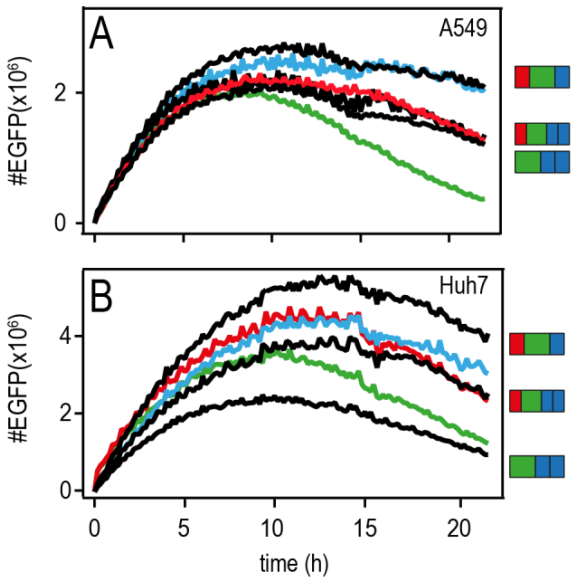

Figure 5 **Mastercurves of the different constructs**

Population averages of A549 (A) and Huh7 (B) cells with the onset time shifted to zero. The green, red and blue curves correspond to the control/5'UTR/3'UTR constructs, respectively. The black curves correspond to the constructs on the right hand side. Intensities were assigned to actual numbers of d2EGFP by calibration as described before (26).

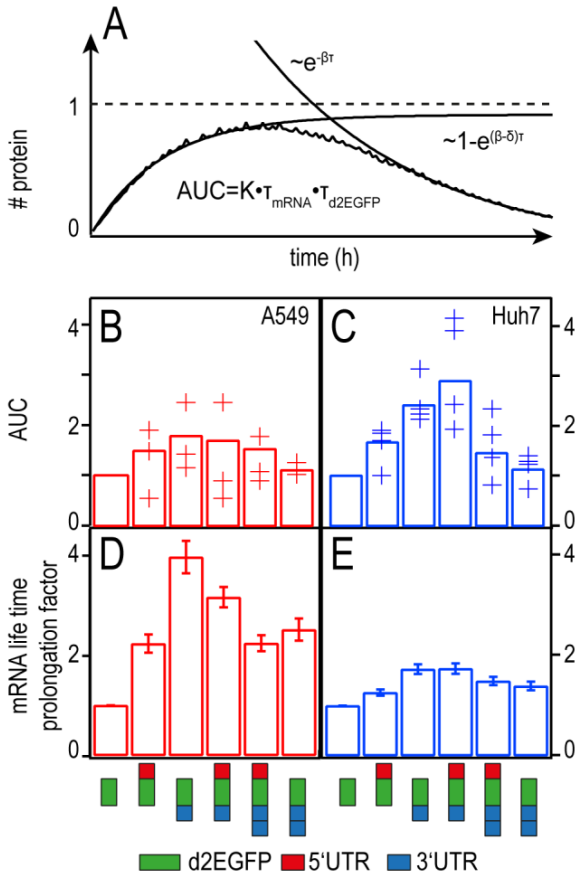

Figure 6 **AUC and mRNA life time prolongation factors of the different constructs**

A) Schematic representation of the AUC to illustrate the interplay between mRNA translation and degradation of mRNA and protein

B) and C) AUC of the different constructs as analysed for $t \rightarrow \infty$. Crosses show relative AUCs of different experiments, the bars correspond to the mean of all single-cell AUCs.

D) and E) mRNA life time prolongation factors. All modifications result in prolonged mRNA life times as compared to the control. Similar trends are observed in A549 (D) and Huh7 (E) cells.

Supplementary Information

Table S1: Secondary structures (mfold)

d2EGFP	∆ G	5'end	3' end	5' UTR	3'UTR
control	-358,9	partial binding with cds (8/8)	loose (8/8)	none	none
5' CYBA	-375	partial binding with 5' CYBA UTR (7/8)	loose (8/8)	binds with cds (6/8)	none
3' CYBA	-411,6	partial binding with 3' CYBA UTR (8/8)	binds with 5'end (4/4)	none	forms one hairpin (7/8)
5'+3' CYBA	-405,7	binds with 3' CYBA UTR (3/8)	binds with 5'end (4/4)	binds with 3' CYBA UTR (4/8)	forms one hairpin (7/8)
5'+2x3' CYBA	-437,7	binds with 3'UTR (8/8)	loose (8/8)	binds with 3'UTR and gene (6/8)	1st 3'UTR: hairpin; 2nd 3'UTR: hairpin (7/8)
2x3' CYBA	-444,1	binds with itself and forms hairpin (8/8)	loose (7/8)	none	1st 3'UTR: hairpin; 2nd 3'UTR: two hairpins (3/8)

In Table S1, features of the mRNA constructs such as free minimum energy (∆G) and secondary structures found at both ends and within the UTRs are listed. The folding platform mfold was used to predict mRNA secondary structures (34). For each construct, we compared the eight secondary structures that have the highest free energy. The highest free energy values are predicted for the 2x3' UTR and the 3' UTR constructs. The 5' end of each mRNA construct partially binds with the 3'UTR or the 5'UTR, except for the control construct, which binds to the coding sequence (cds). Interestingly, the 5' end of the 2x3' mRNA construct forms a stabilizing hairpin with itself. However, hairpin loops near 5'end can also hinder protein translation (33). Another feature was found in the 3' end of the 3' UTR and 5'+3' UTR mRNA constructs: There, the 3' end binds with the 5' end, minimizing the distance from each other and thus enabling faster initiation of translation. Unlike the 5'UTRs, the 3' UTR of each mRNA construct forms at least one hairpin with itself.

Table S2: Human CYBA and its UTRs

Untranslated region	DNA sequence (from 5' to 3')
5'	CGCGCCTAGCAGTGTCCCAGCCGGGTTCGTGTCGCC
3'	CCTCGCCCCGGACCTGCCCTCCCGCCAGGTGCACCCACCTGC**AATAAA**TGCAGCGAAGCCGGGA

Table S2 shows the exact genetic code of the human CYBA gene UTRs. DNA sequences are shown from the 5' to the 3' end. The poly-adenylation signal (PAS) of the 3' UTR is shown in bold letters and the insulin 3'UTR stability element (INS_SCE) is underlined. The 5' UTR consists of 36 base pairs, whereas the 3' UTR contains 64 base pairs. Both UTRs are shorter than average human UTRs, which consist of around 200 nucleotides in the case of 5'UTRs and approximately 1000 nucleotides in the case of 3'UTRs.

Transfection efficiencies on microstructured substrates

The percentage of successfully transfected cells was assessed to compare two different transfection agents and to ensure that transfection efficiencies were not hampered by microstructured cell growth (see Figure S1). Here, all cells grew on microstructured protein arrays. We obtained higher transfection efficiencies for Lipofectamine^{TM}2000 as compared to DOGTOR. Using a commercial Live/Dead cell viability assay (Molecular Probes, Germany), we found high cell viability rates above 80% (data not shown).

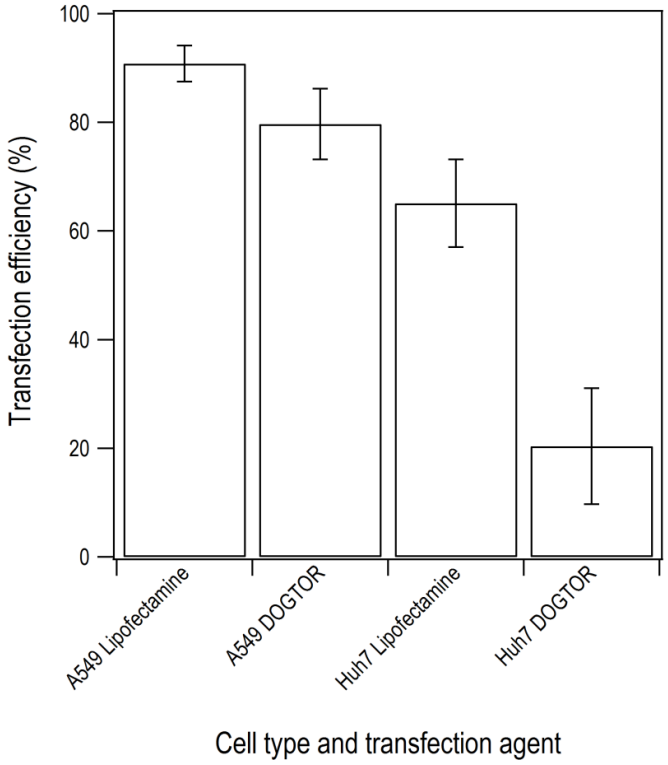

Figure S1 **Transfection efficiencies on microstructured substrates**

Percentage of transfected cells and corresponding standard deviations for A549 cells and Huh7 cells transfected with SNIM RNA with help of Lipofectamine^{TM}2000 or DOGTOR. We find higher transfection efficiencies for cells transfected with Lipofectamine^{TM}2000.

Area under the curve (AUC)

Assuming biochemical rate equations for translation and degradation, the amount of expressed protein after mRNA transfection can be written as

$G_{mRNA}(t) = \frac{K}{\delta - \beta} \cdot \left(1 - e^{-(\delta-\beta)(t-t_0)}\right) \cdot e^{-\beta(t-t_0)}$ (Equation 1 of the main text).

The cumulative efficacy of an mRNA, known as the AUC, can then be calculated for long times ($t \to \infty$) as follows:

$$AUC = \int_{t=0}^{t=\infty} G(t) = \frac{K}{\delta - \beta} \int_0^{\infty} \left[e^{-\beta\tau} - e^{-\delta\tau}\right] d\tau = \cdots = \frac{K}{\delta \cdot \beta} \text{ with } \tau = t - t_0.$$

This can be rewritten as

$AUC = m_0 \cdot k_{TL} \cdot \tau_{mRNA} \cdot \tau_{d2EGFP}\ with\ m_0 \cdot k_{TL} = K, \tau_{mRNA} = \frac{ln2}{\delta}, and\ \tau_{d2EGFP} = \frac{ln2}{\beta}.$

Hence, the expression of a therapeutic protein after mRNA transfection can be interpreted as a trade-off between translational efficiency and degradation as illustrated in Figure 3 C of the main text.

The single-cell AUC distributions can be seen in Figure S2. Because the AUC depends linearly from the mRNA and protein life times, the single-cell AUC distributions are closely related to the mRNA and protein half-life distributions that are shown in Figure 4B and 4C of the main text.

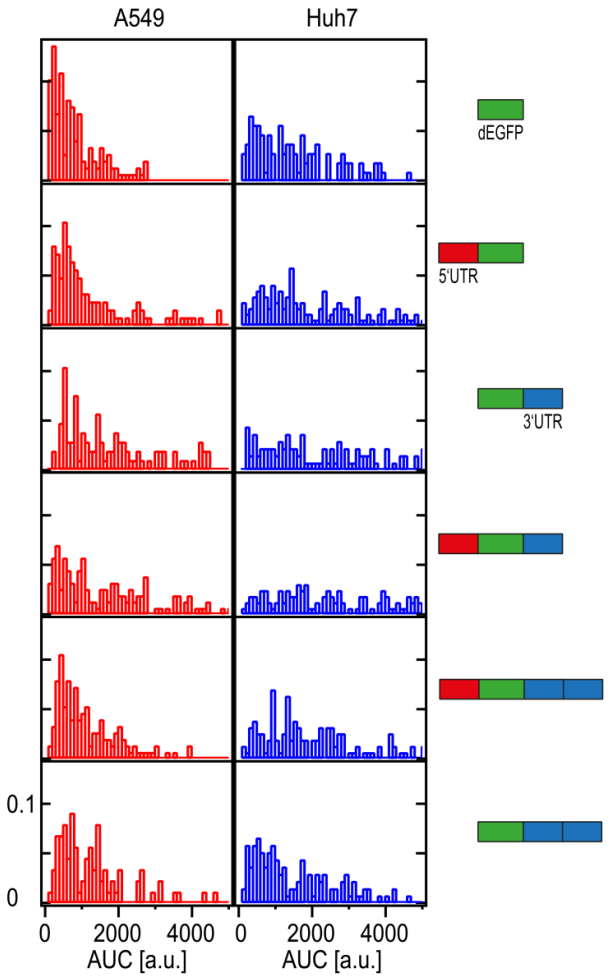

Figure S2 **Distribution of the single-cell AUCs**

Degradation rate of the reporter protein

To check the fitted d2EGFP degradation rates, we independently measured the degradation rate of d2EGFP inside A549 and Huh7 cells in microstructured six-channel slides. Protein synthesis was blocked by the antibiotic cycloheximide, which interferes with peptidyl transferase activity (35). Single-cell fluorescence intensity time courses were monitored for approximately 20 hours (see Figure S3). Control experiments ensured that the decrease in fluorescence intensity was not due to photobleaching of the chromophore. Single-cell time courses were fitted by a single exponential fit, yielding distributions of protein degradation rates. The mean degradation rates were found to be 0.28/h (std 0.08/h) in A549 cells and 0.17/h (std 0.08/h) in Huh 7 cells, corresponding to protein life times of 2.46 hours and 4.04 hours, respectively. Although these life times are significantly shorter than the life times as determined by single-cell time course analysis of mRNA mediated protein expression, the ratio between the mean life times of d2EGFP inside Huh7 and A549 cells is the same (4.04h/2.46h=1.64 as measured by translational blocking as compared to 7.4h/4.5h=1.64 as determined by fitting the analytical solution for mRNA expression).

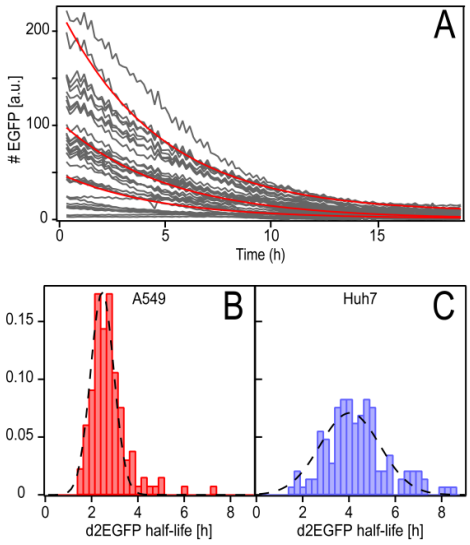

Figure S3 **Distributions of directly measured d2EGFP half-lives**

A) Exemplary time courses of cycloheximide-induced d2EGFP degradation in Huh7 cells. Red lines are simple exponential fits for protein degradation.

B) Distribution of d2EGFP half-lives measured in A549 cells, yielding a mean half-life of 2.46 hours (std 0.71 h).

C) Distribution of d2EGFP half-lives measured in Huh7 cells, yielding a mean half-life of 4.04 hours (std 1.82 h).

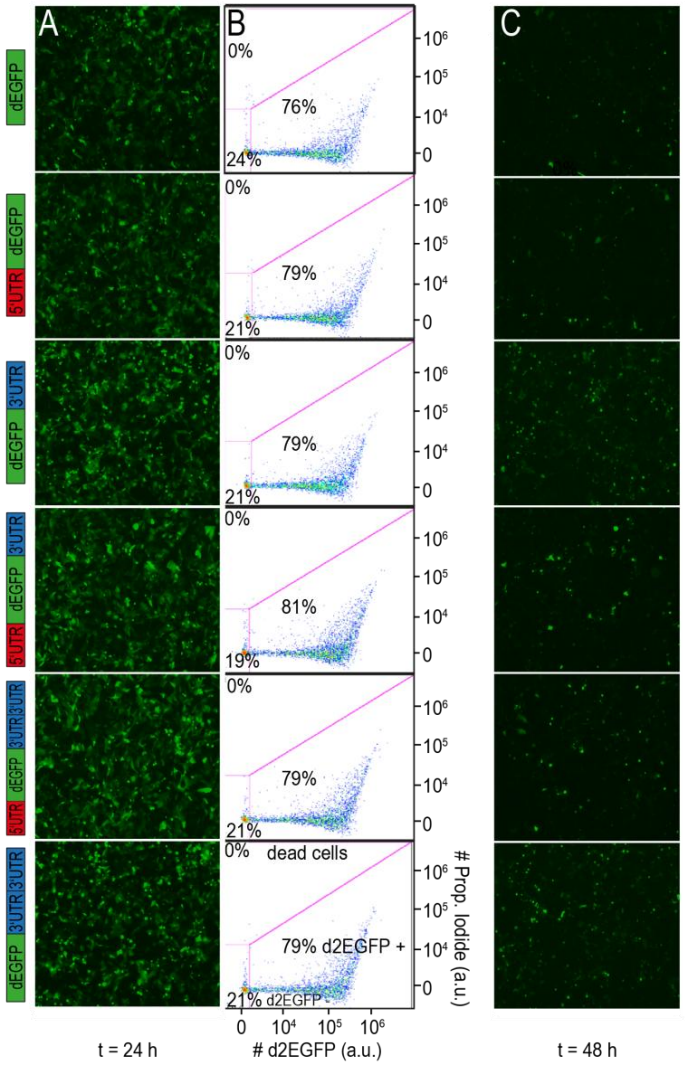

Figure S4 **Fluorescence microscopy and flow cytometry data of Huh7 cells**

A) Fluorescence microscopy pictures taken with 4-fold magnification (JULY^{TM}) at 24 hours post-transfection. All constructs showed improved protein expression levels as compared to the control.

B) The percentage of d2EGFP positive cells as determined by FC is similar for all constructs. Propidium iodide was used to detect dead cells. The applied gates ensured exclusion of dead cells and untransfected cells.

C) At 48 hours post transfection, sustained protein expression was higher for the stabilized constructs as compared to the control.

References

1. Tavernier G*, et al.* (2011) mRNA as gene therapeutic: how to control protein expression. *J Control Release* **150**(3):238-247.
2. Yamamoto A, Kormann M, Rosenecker J, & Rudolph C (2009) Current prospects for mRNA gene delivery. *Eur J Pharm Biopharm* 71(3):484-489.
3. Kormann MSD*, et al.* (2011) Expression of therapeutic proteins after delivery of chemically modified mRNA in mice. *Nat Biotech* 29(2):154-157.
4. Esteller M (2011) Non-coding RNAs in human disease. *Nat Rev Genet* 12(12):861-874.
5. Kariko K & Weissman D (2007) Naturally occurring nucleoside modifications suppress the immunostimulatory activity of RNA: implication for therapeutic RNA development. *Current Opinion in Drug Discovery and Development* 10(5):523.
6. Pesole G, Grillo G, Larizza A, & Liuni S (2000) The untranslated regions of eukaryotic mRNAs: Structure, function, evolution and bioinformatic tools for their analysis. *Briefings in Bioinformatics* 1(3):236-249.
7. Pestova TV, Lorsch JR, & Hellen CU (2007) 4 The Mechanism of Translation Initiation in Eukaryotes. *Cold Spring Harbor Monograph Archive* 48:87-128.
8. Barrett L, Fletcher S, & Wilton S (2012) Regulation of eukaryotic gene expression by the untranslated gene regions and other non-coding elements. *Cell. Mol. Life Sci.* 69(21):3613-3634.
9. Mignone F*, et al.* (2005) UTRdb and UTRsite: a collection of sequences and regulatory motifs of the untranslated regions of eukaryotic mRNAs. *Nucleic Acids Research* 33(suppl 1):D141-D146.
10. Moore MJ (2005) From Birth to Death: The Complex Lives of Eukaryotic mRNAs. *Science* 309(5740):1514-1518.
11. Pichon X*, et al.* (2012) RNA Binding Protein/RNA Element Interactions and the Control of Translation. *Current Protein & Peptide Science* 13(4):294-304.
12. 't Hoen PAC*, et al.* (2011) mRNA degradation controls differentiation state-dependent differences in transcript and splice variant abundance. *Nucleic Acids Research* 39(2):556-566.
13. Pesole G*, et al.* (2001) Structural and functional features of eukaryotic mRNA untranslated regions. *Gene* 276(1–2):73-81.
14. Gebauer F & Hentze MW (2004) Molecular mechanisms of translational control. *Nat Rev Mol Cell Biol* 5(10):827-835.
15. Tillmar L, Carlsson C, & Welsh N (2002) Control of Insulin mRNA Stability in Rat Pancreatic Islets: REGULATORY ROLE OF A 3′-UNTRANSLATED REGION PYRIMIDINE-RICH SEQUENCE. *Journal of Biological Chemistry* 277(2):1099-1106.
16. Holtkamp S*, et al.* (2006) Modification of antigen-encoding RNA increases stability, translational efficacy, and T-cell stimulatory capacity of dendritic cells. *Blood* 108(13):4009-4017.
17. Moor CHd, Meijer H, & Lissenden S (2005) Mechanisms of translational control by the 3′ UTR in development and differentiation. *Seminars in Cell & Developmental Biology* 16(1):49-58.
18. Garneau NL, Wilusz J, & Wilusz CJ (2007) The highways and byways of mRNA decay. *Nat Rev Mol Cell Biol* 8(2):113-126.
19. Szostak E & Gebauer F (2012) Translational control by 3′-UTR-binding proteins. *Briefings in Functional Genomics*.
20. Tyagi S (2009) Imaging intracellular RNA distribution and dynamics in living cells. *Nat Meth* 6(5):331-338.
21. Park HY*, et al.* (2014) Visualization of Dynamics of Single Endogenous mRNA Labeled in Live Mouse. *Science* 343(6169):422-424.
22. Miller C*, et al.* (2011) Dynamic transcriptome analysis measures rates of mRNA synthesis and decay in yeast. *Molecular Systems Biology* 7(1).
23. Nolan T, Hands RE, & Bustin SA (2006) Quantification of mRNA using real-time RT-PCR. *Nat. Protocols* 1(3):1559-1582.

24. Rabani M*, et al.* (2011) Metabolic labeling of RNA uncovers principles of RNA production and degradation dynamics in mammalian cells. *Nat Biotech* 29(5):436-442.
25. Schwanhausser B*, et al.* (2011) Global quantification of mammalian gene expression control. *Nature* 473(7347):337-342.
26. Leonhardt C*, et al.* (2014) Single-cell mRNA transfection studies: Delivery, kinetics and statistics by numbers. *Nanomedicine: Nanotechnology, Biology and Medicine* 10(4):679-688.
27. Rottgermann PJF, Alberola AP, & Radler JO (2014) Cellular self-organization on micro-structured surfaces. *Soft Matter*.
28. Corish P & Tyler-Smith C (1999) Attenuation of green fluorescent protein half-life in mammalian cells. *Protein Engineering* 12(12):1035-1040.
29. Halter M, Tona A, Bhadriraju K, Plant AL, & Elliott JT (2007) Automated live cell imaging of green fluorescent protein degradation in individual fibroblasts. *Cytometry Part A* 71A(10):827-834.
30. Li X*, et al.* (1998) Generation of Destabilized Green Fluorescent Protein as a Transcription Reporter. *Journal of Biological Chemistry* 273(52):34970-34975.
31. Chen J-M, Férec C, & Cooper D (2006) A systematic analysis of disease-associated variants in the 3′ regulatory regions of human protein-coding genes II: the importance of mRNA secondary structure in assessing the functionality of 3′ UTR variants. *Hum Genet* 120(3):301-333.
32. Gaspar P, Moura G, Santos MA, & Oliveira JL (2013) mRNA secondary structure optimization using a correlated stem–loop prediction. *Nucleic acids research* 41(6):e73-e73.
33. Babendure JR, Babendure JL, Ding J-H, & Tsien RY (2006) Control of mammalian translation by mRNA structure near caps. *RNA (New York, N.Y* 12(5):851-861.
34. Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic acids research* 31(13):3406-3415.
35. Siegel MR & Sisler HD (1963) Inhibition of Protein Synthesis in vitro by Cycloheximide. *Nature* 200(4907):675-676.

5 Predictive modeling of mRNA lipofection

Predictive modeling is highly likely to take its place in future mainstream biological research. The integration and evaluation of quantitative experimental data and the subsequent integration of such information into the model allow the generation of predictions and a focusing of experimental directions. Figure 5.1 illustrates the interplay of experimental data and theory in the context of predictive modeling approaches.

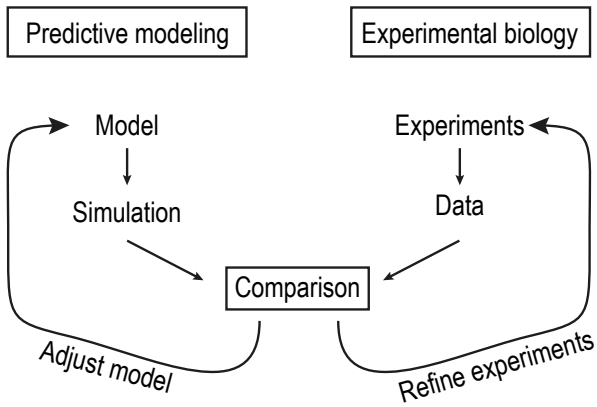

Figure 5.1: *Methodology of predictive modeling. Quantitative data are provided by experimental biology and compared to the outcome of simulation runs of the model to test hypotheses as well as to optimise and refine both the model and the experiments' design. (Adapted by permission from Macmillan Publishers Ltd: Nature Biotechnology [106], copyright 2007)*

The following publication P2 reports on this multi-level kinetic model, which covers the cooperative effects of vectors during cellular uptake. Because the predictive power of a model highly relies on experimentally substantiated, quantitative data, we used continuous, long-term quantitative time-lapse microscopy and single-cell time course analyis as a basis. This mRNA transfection model is an extension of the model that was introduced in P1 [54]. In the refined version, the lipoplexes are compartmentalised and varying numbers of lipoplexes are allowed per endosome. However, lipoplex unpacking and degradation inside endosomes is not covered by the model but proposed for a future model refinement. Deterministic simulations are compared with stochastic Monte Carlo simulation on the basis of the Gillespie algorithm, which was a first stochastic formulation of chemical kinetics [156]. Our model covers the multi-level nature of the lipoplex pathway in mRNA lipofection. Unlike existing models, it explicitly reflects the compartments that are involved in the transfer

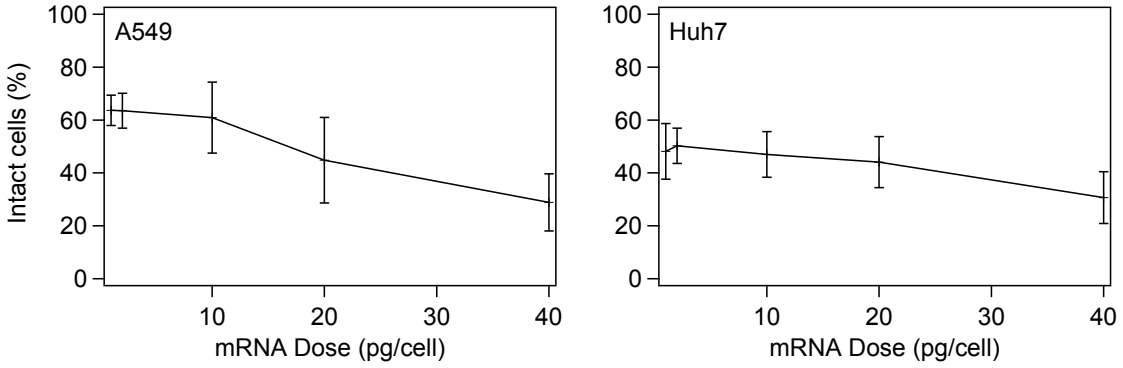

Figure 5.2: *Dose-dependent cell viability rates of A 549 and Huh 7 cells. Cell viability, as reflected by the percentage of intact cells, suffers from increasing amounts of the cationic helper lipid and toxic effects caused by protein-overexpression. Mean values and standard deviations of triplicates (n=10000) are shown.*

process. Our approach was to generate simulated time courses of protein expression on the basis of the model, which comprises the delivery as well as the expression phase of mRNA transfection. Thereby obtained *in silico*-data were treated in exactly the same way as experimental data, *i.e.* resulting time courses were fitted using the analytical solution of the biochemical rate equations as described in [54]. Subsequently, parameter optimisation was done on the basis of the resulting distributions of both onset times and expression levels, as well as the simulated dose-response curve. A simulated annealing algorithm, which is a standard technique for parameter optimisation, was used. The name of this global optimisation method is inspired by the slow cooling processes in metallurgy, termed annealing. In the algorithm, this slow temperature decrease is reflected by a slow decrease of the probability of accepting inferior solutions while searching the parameter space. The fact that such worse solutions are occasionally accepted by the algorithm ensures that the entire solution space is sampled, yielding global optima. Hence, the simulated annealing algorithm is not gradient-based nor influenced by the initial parameter guess and thus well-suited for parameter optimisation of stochastic simulations like ours. We employed known rate-constants to reduce the number of free parameters. The remaining five kinetic rates were then fitted and optimal agreement of simulated and experimental data was achieved.

It should be mentioned that toxic effects with increasing mRNA dose, which can be seen in figure 5.2, are not covered by our model or other existing models. The percentage of intact cells decreases with increasing amounts of expressed protein and helper lipids. An excess of lipoplexes significantly compromises cell viability, which for example has been observed for pDNA lipoplexes in a different type of cells [157]. As a consequence, the direct comparison of experimental and simulated transfection efficiencies at high doses is rather delicate. Also, our model can only release whole packages of mRNA, which most likely is not a realistic scenario. However, regardless of the chosen model, some parameters inherently have to be approximated. Our study demonstrates the power and limits of existing toolboxes for biological modeling and their potential to ultimately enhance the communication between experimentalists and theoreticians.

A future, refined model that allows lipoplex unpacking and degradation of mRNA inside endosomes could be used to test hypotheses such as the assumption that longer mRNA life times result in more mRNA molecules surviving the endosomal phase and hence increasing transfection efficiencies and protein expression levels. For corresponding experiments, differently stabilised mRNA constructs as used for the study of chapter 4 could be employed.

The results of this project are presented in the following publication P2.

5.1 Associated publication P2

Multi-Level Kinetic Model of mRNA Delivery via Transfection of Lipoplexes

By

Thomas S. Ligon, Carolin Leonhardt, Joachim O. Rädler

published in
PLoS ONE 9(9): e107148, 2014.
doi: 10.1371/journal.pone.0107148

OPEN ACCESS Freely available online

PLOS ONE

Multi-Level Kinetic Model of mRNA Delivery via Transfection of Lipoplexes

Thomas S. Ligon*, Carolin Leonhardt, Joachim O. Rädler

Faculty of Physics and Center for NanoScience (CeNS), Ludwig-Maximilians-Universität, München, Germany

Abstract

Recent work on the use of mRNA lipoplexes for gene delivery demonstrates the need for a mathematical model that simulates and predicts kinetics and transfection efficiency. The small copy numbers involved make it necessary to use stochastic models and include statistical analysis of the variation observed in the experimental data. The modeling requirements are further complicated by the multi-level nature of the problem, where mRNA molecules are contained in lipoplexes, which are in turn contained in endosomes, where each of these entities displays a behavior of its own. We have created a mathematical model that reproduces both the time courses and the statistical variance observed in recent experiments using single-cell tracking of GFP expression after transfection. By applying a few key simplifications and assumptions, we have limited the number of free parameters to five, which we optimize to match five experimental determinants by means of a simulated annealing algorithm. The models demonstrate the need for modeling of nested species in order to reproduce the shape of the dose-response and expression-level curves.

Citation: Ligon TS, Leonhardt C, Rädler JO (2014) Multi-Level Kinetic Model of mRNA Delivery via Transfection of Lipoplexes. PLoS ONE 9(9): e107148. doi:10.1371/journal.pone.0107148

Editor: Roeland M.H. Merks, Centrum Wiskunde & Informatica (CWI) & Netherlands Institute for Systems Biology, Netherlands

Received March 22, 2014; **Accepted** August 14, 2014; **Published** September 19, 2014

Data Availability: The authors confirm that all data underlying the findings are fully available without restriction. All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by the Deutsche Forschungsgemeinschaft (DFG) via project B1 within the SFB 1032, the Excellence Cluster 'Nanosystems Initiative Munich (NIM)', and FP7 EU grants NanoTransKinetics and Nano-MILE. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* Email: Thomas.Ligon@physik.uni-muenchen.de

Introduction

Quantitative analysis of transfection is important for gene therapy involving plasmid DNA and mRNA, as well as high-throughput screening (HTS) and siRNA research [1–4]. For this reason, it is important to know more about the kinetics and dose-response relationship for delivery of genes and RNA-based nucleic acid constructs and to understand the common principles that underlie nucleic acid pharmacokinetics in any given cell type. Many studies have collected quantitative data on the uptake and pathway of gene carriers [5–10] and the physico-chemical characterization of cationic lipoplexes and polyplexes has been reviewed extensively [11–17]. In the last few years, first theoretical considerations modeling the uptake and pharmacokinetics of lipolexes using biochemical reaction kinetics have been undertaken [18–20]. Some specialized models also address the spatial distribution and active transport along microtubules [21]. The stochastic nature of in the delivery process has been investigated for nanoparticles [22] and for plasmid DNA [23]. The use of movies for the analysis of single-cell tracking experiments has been reviewed [24]. For modelling of biological systems in general, there is an emerging set of tools in the context of systems biology, including a new generation of computational methods, such as process calculi and "executable biology" [25]. In fact, many biological reactions require addition of stochastic modeling as well as spatial aspects that go beyond reaction and diffusion [26]. For example, endosomes contain lipoplexes and lipoplexes contain mRNA molecules, and this can lead to a combinatorial explosion in the number of variables and equations. The transfection process requires the use of modeling techniques that have not been used often, because substances can be contained in each other.

The problem of multi-level modeling has been treated in many investigations and tools. Systems Biology Markup Language (SBML) [27] and tools based on it, for example Cell Designer [28] and Copasi [29], include the concept of compartments, which contain species, but the compartments are only containers that cannot support reactions of their own. First attempts to allow modelling with compartments include the process calculus Pi Calculus [30–35] and tools based on it, such as BioAmbients [36], Beta-Binders [37–39] and the Stochastic Pi Machine SPiM [40]. In addition, the "rules-based" language BioNetGen Language BNGL [41] and tools based on it, such as NFsim [42], contain some very explicit methods for handling nested structures. One example where these techniques were used is a model for the uptake of nanoparticles is the work by Dobay et al. using SPiM [43], which also demonstrates the need for multi-level modeling in many situations involving nanoparticles.

Recently, we showed that quantitative analysis of transfection at the single-cell level makes it possible to analyze the stochastic aspects of transfection quantitatively [23,44]. The single cell exhibits time courses that are characterized by a distinct delay time before the onset of expression, a phase of GFP increase and finally a steady state level. We showed that the distribution of steady-state levels was related to the number of successfully delivered plasmids and well described by an analytical model [23]. In the same spirit, we analyzed the transfection of mRNA, which is more homogeneous

and earlier compared to pDNA [45,46]. However, there is yet little understanding regarding the kinetics of mRNA delivery. It is generally accepted that mRNA lipoplexes are taken up via clathrin-dependent endocytosis [47]. Existing models for RNA delivery sometimes include a single "internalization" reaction, but that is not sufficient for reproducing the data created by single-cell tracking experiments. In particular, there is no kinetic model for the delivery of mRNA that explicitly takes the compartments of the transfer process into account.

Here we present a mathematical model, based on mass-action kinetics, which describes the uptake of mRNA lipoplexes via endocytosis and endosomal lysis. Our goal was to create a kinetic model that reproduces experimental data, especially the distribution of time courses, and supports predictive modeling. While the investigation of plasmid DNA [23] provides some background and motivation, this model was based solely on the data published on the experiment with mRNA [44]. We demonstrate that the uptake kinetics is well described by a stochastic, mass action based model that accounts for uptake of multiple lipoplexes. We solve the problem of parameter estimation by choosing well-known rate constants from literature and keeping five kinetic rates free, which we optimize to meet the constraints of the experimental transfection statistics and measured onset time distribution by using a simulated annealing algorithm. As such, the model yields uptake behavior that reproduces the experimental data and is capable of predicting behavior beyond the experimental parameter regimes. The model also demonstrates the need for modeling of nested species as well as modeling kinetic reactions in a stochastic version in order to reproduce the shape of the dose-response and expression-level curves, and the need to include the maturation step in order to reproduce the variance of the onset-time distribution. The benefit of predictive modeling and the known limitations of the model are discussed.

Model Description

Streamlined Model

We model mRNA transfection by a sequence of mass-action type chemical reactions (shown in Figure 1), which can be divided into the delivery of lipoplexes and the GFP expression via the mRNA released.

The delivery phase is described by the following ODEs:

$$\frac{dL_{ex}}{dt} = -k_A L_{ex} - k_W L_{ex} \quad (1)$$

$$\frac{dP}{dt} = +k_A L_{ex} - k_E P \quad (2)$$

$$\frac{dE}{dt} = +k_E P - k_L E - d_E E \quad (3)$$

$$\frac{dL_{in}}{dt} = +k_L E - k_U L - d_L L \quad (4)$$

$$\frac{dM}{dt} = +350 k_U L_{in} - d_M M \quad (5)$$

Where L_{ex} is the concentration of external lipoplexes, k_A is the rate at which lipoplexes attach to the cell surface, k_W is the washing rate, which is equal to zero at first and jumps to a high value after the incubation time or normally one hour, P is the concentration of clathrin-coated pits (i.e. number per cell), k_E is the rate of endocytosis, E is the concentration of endosomes (i.e. number per cell), k_L is the rate of lysis of endosomes, d_E is the rate of endosome degradation, L_{in} is the concentration of internal lipoplexes, k_U is the rate of lipoplex unpacking, d_L is the rate of degradation of lipoplexes, M is the concentration of mRNA, k_U is the rate of unpacking of lipoplexes, and d_M is the rate of degradation of mRNA. The degradation of endosomes is primarily a model parameter to represent endosomes that are never observed to lyse, and includes mRNA degradation in the endosome.

The expression phase is described by the following ODEs, plus equation (5), which includes mRNA degradation:

$$\frac{dG}{dt} = +k_{TL} M - k_M G - d_G G \quad (6)$$

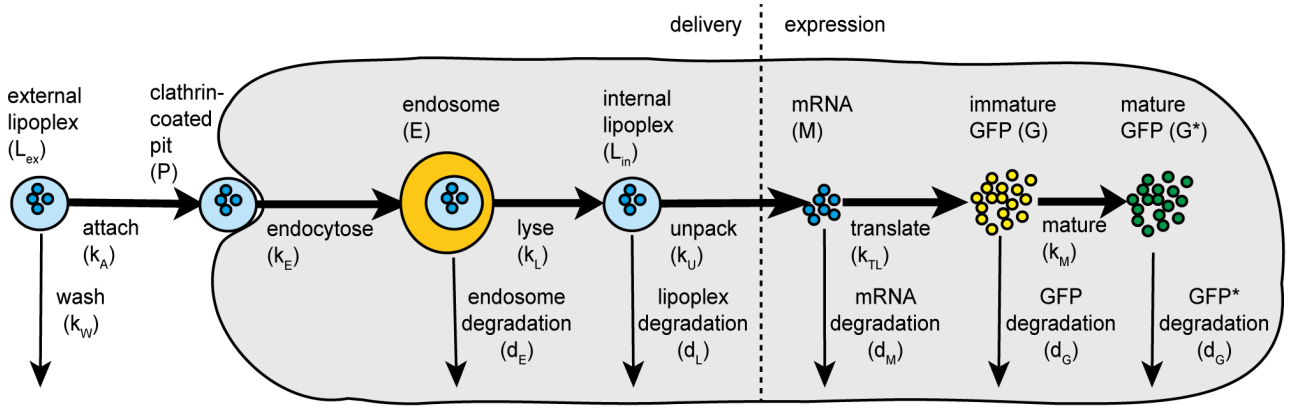

Figure 1. Diagram of the streamlined transfection model. External (extracellular) lipoplexes attach to the surface of the cell, forming clathrin-coated pits, which enter the cell via endocytosis, leading to the formation of endosomes, which either lyse or degrade. This puts the lipoplexes into the cytosol, where they unpack, releasing the mRNA, which translates to unfolded GFP molecules, which then mature (folding and oxidation), to produce active GFP. In addition to the endosomes, the lipoplexes, mRNA, immature and mature GFP are all degraded at set rates.
doi:10.1371/journal.pone.0107148.g001

$$\frac{dG^*}{dt} = +k_M G - d_G G^* \quad (7)$$

Where G is the concentration of immature (unfolded) GFP, k_{TL} is the rate of translation, k_M is the rate of maturation (folding and oxidation), d_G is the rate of degradation of both immature and mature GFP, and G^* is the concentration of mature GFP. The reaction rates are documented in Table 1.

This first model shows a very linear progression of single lipoplexes attaching to and entering the cell, but we know from experiment that endosomes can contain multiple lipoplexes, so we need to address that and allow for endocytosis of multiple lipoplexes per endosome. This means that we will have multiple levels of containment.

Multi-Level Modeling

The solution to the complexity that arises from multiple levels of structure is a key aspect of the model shown in Figure 2, so we will describe it here in very general terms. For readers who are interested in more detail, the File S1 contains the code of all versions of the model.

The initial condition of external lipoplexes provides a first example of this. In ordinary differential equations, we would use the name of the lipoplexes (Lext or L_{ex}) as a variable in the equations. This variable refers to the concentration of lipoplexes, or, equivalently, the number of particles in a given volume. In an SBML-based [27] tool, this is also called a species. Now the problem here is that the lipoplexes come in different sizes, based on the number of mRNA molecules they contain. In the current experimental situation we are modeling, the lipoplexes have a mean diameter of 120 nm and a standard deviation of 10 nm. This size was determined by fluorescence correlation spectroscopy (data not shown). When we additionally take the packing density of the lipoplexes into account, this size corresponds to a mean of 350 mRNA molecules per lipoplex and lipoplex sizes ranging from 270 to 445 mRNA molecules. See Supplementary data of Leonhardt et al. [44] for a detailed description.

There are three solutions to this problem. First, we can use a tool in which we can include a parameter for the size of the lipoplex. In other words, we can write Lext(n), where n is the number of mRNA molecules, and use that in the model. Second, as an alternative, we can simply list all possible values of the size as separate species, e.g. Lext270, Lext271 … Lext445. Finally, we can apply a key simplification and assume that all lipoplexes contain exactly 350 mRNA molecules.

Next, we need to consider the endosomes. Our experience with both experimental data and modeling shows us that each endosome can only contain a small number of lipoplexes, and we are safe when we set this to an arbitrary maximum of 10. In addition, each of those lipoplexes can contain anywhere from 270 to 445 mRNA molecules. In order to list all of these cases, we would need more than 175^{10} different variables (or species), something that is clearly impossible.

The key simplification in this paper, assuming that all lipoplexes have the same size, along with listing all possible endosome sizes, makes it possible to formulate the model in SBML and use Copasi to run the simulations. We have also evaluated the use of other tools and present those results here, for the benefit of experts in those tools and modeling techniques in general. The second implementation uses Pi-Calculus-based SPiM and preserves full

Table 1. Rates.

A parameters, fitted (optimized) and fixed						
	role	goal (exp.)	streamlined with slow maturation	multiple-lipoplex with fast maturation	multiple-lipoplex with slow maturation	literature
k_A (attach)	fitted		.03	0.26	0.27	0.006–0.5 [20,21,59]
k_E (endocytosis)	fitted		.8	0.73	0.81	0.16–0.5 [20,21,59]
k_L (lysis)	fitted		.065	0.10	0.11	0.001–0.96 [20,21,59]
k_M (maturation)	fitted or fixed		5.5	9.23	5.5	0.5–9.23 [48,60–63]
d_E (endosome degradation)	fitted		0.65	0.60	0.67	n.a.
k_U (unpack)	fixed		1e+06	1e+06	1e+06	n.a.
d_L (lipoplex degradation)	fixed		1e–06	1e–06	1e–06	n.a.
k_{TL} (translation)	fixed		170	170	170	170 [44]
d_M (mRNA degradation)	fixed		0.062	0.062	0.062	0.062 [44]
d_G (GFP degradation)	fixed		0.056	0.056	0.056	0.056 [44]
B experiment vs. simulation						
TE (transfection efficiency)	target	40	44	36	38	
LC (lipoplexes on cell)	target	6	6.43	6.02	6.03	
maxGFP	target	7.09e+5	4.32e+5	4.91e+5	5.34e+5	
t0-mean	target	3.14	3.36	3.49	3.23	
t0-width	target	1.54	1.72	2.05	1.65	

A) The table shows the rate constants used by the simulation. During optimization, k_A, k_E, k_L, and d_E were varied, and k_M was varied in one case. Column "streamlined with slow maturation" is the streamlined model with $k_M = 5.5$ fixed. Column "multiple-lipoplex with fast maturation" is the multiple-lipoplex model with $k_M = 9.23$ fixed to the value from literature. Column "multiple-lipoplex with slow maturation" is the multiple-lipoplex model with k_M varied (optimized). The literature values are described in more detail in the File S1. B) The last 5 rows are the experimental data used as a goal in optimization.

doi:10.1371/journal.pone.0107148.t001

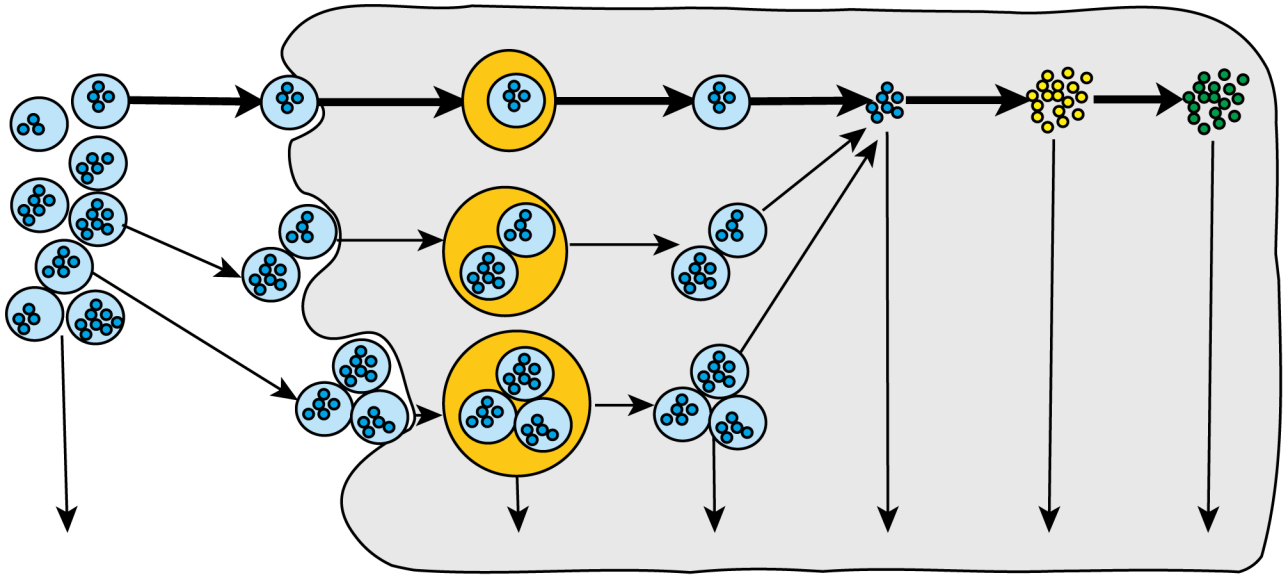

Figure 2. Diagram of the multiple-lipoplex transfection model. This includes the same processes as in the streamlined model, except that here the clathrin-coated pits and the endosomes can contain multiple lipoplexes.
doi:10.1371/journal.pone.0107148.g002

complexity, except that we used a smaller width for the lipoplex size distribution in order to keep the code smaller. The variable sizes of the lipoplexes are kept throughout their lifetime, and the variable sizes of pits and endosomes are represented by listing all possible values, due to limitations in formulating reactions of parameters in SPiM (as opposed to processes). The third version uses the rule-based language BioNetGen Language (BNGL) in the tool NFsim, and exposes a limitation that prevents us from using a parameter (such as the number of mRNA molecules in a lipoplex) in a reaction without setting it to an explicit value.

Multiple-Lipoplex Model

The multiple-lipoplex model (Figure 2) follows the lines of the streamlined model (heavy arrows), but also includes the formation of clathrin-coated pits that include multiple lipoplexes.

The delivery phase is described by the following ODEs:

$$\frac{dL_{ex}}{dt} = -k_{AX}L_{ex} - \sum_{i=1}^{9} k_{AX}L_{ex}P_i - k_W L_{ex} \tag{8}$$

$$\frac{dP_i}{dt} = +k_{AX}L_{ex} - k_{AX}L_{ex}P_i - k_E P_i \tag{9}$$

$$\frac{dP_{i+1}}{dt} = +k_{AX}L_{ex}P_i - k_{AX}L_{ex}P_{i+1} - k_E P_{i+1} \qquad i=1...9 \tag{10}$$

$$\frac{dE_i}{dt} = +k_E P_i - k_L E_i - d_E E_i \qquad i=1...10 \tag{11}$$

$$\frac{dL_{in}}{dt} = +\sum_{i=1}^{10} ik_L E_i - k_U L_{in} - d_L L_{in} \tag{12}$$

and equation (5) from above, where P_i is the concentration of clathrin-coated pits of size i, i.e. containing i lipoplexes, E_i is the concentration of endosomes of size i, and the new rate of attachment is k_{AX} calculated by dividing k_A by the number of pits plus one, in order to assure a constant rate of attachment even when the number of pits increases. All other symbols are the same as in the streamlined model.

The expression phase is described by the same ODEs as in the streamlined model, (5), (6), and (7).

This model, in contrast to the streamlined model, includes different-sized lipoplexes, with their sizes preserved through all reactions up to unpacking. This seemingly easy extension allowing variable lipoplex sizes and variable endosome sizes leads to a severe combinatorial explosion of species and reactions. For the analysis included in this paper, we have avoided a large part of this issue by assuming that all lipoplexes have the same size. This is a very significant simplification, but nevertheless allows fairly good simulation results, and makes it possible to run simulations both deterministically and stochastically, and also to run parameter estimation.

We created 3 implementations of the model. The first is written in SBML, was run in Copasi, and assumes a very significant simplification (all liposomes have the same size); it was used for the analysis in this paper. The second is written in Pi Calculus and was run in the Stochastic Pi Machine (SPiM), and includes a limited example of variable-sized lipoplexes. The third is written in BNGL and was tested in NFSim.

Parameter Optimization

In order to compare the model to the experimental data, the best values need to be found for the five parameters that have been left free, such as the rate of endocytosis. This requires adjusting the model to best fit the five experimental determinants, such as the dose-response relationship. However, since the experimental data

is based on single-cell tracking, it includes the variance of the distributions of multiple time courses. As a result, each attempt to find a better value for the parameters requires two steps: First, it is necessary to run the simulation many times (typically 1,000–5,000) and second, to compare the distributions with the experimental data. In all cases where we compare simulation data to experimental data, we use the same analytical model for the expression phase and the same fitting procedures for both data sets, in order to make a good comparison between simulation and experiment, as reported in [44].

Since we are optimizing a stochastic model, we have chosen to use the simulated annealing algorithm. This algorithm chooses a new set of values for the parameters, based on random numbers, then runs the two steps of simulation and analysis described above, and compares the results with the experimental data. The comparison involves the current value of a "temperature" variable and the Boltzmann function in order to allow the algorithm to move away from local optima that may not be globally optimal.

The first two parameters in the model are the initial concentration of external lipoplexes and the incubation time (time until the cells are washed). These parameters are not part of the optimization process, since they are determined by experiment, but they do appear in the plots we have created of the dose-response relationship and incubation dependency, which we also compare with experimental data. In addition, we have varied these parameters as part of predictive modeling.

The parameters in the optimization process are the rates of attachment, endocytosis, and lysis, along with the rate of endosome degradation, plus the rate of GFP maturation. We optimize these five parameters to match five data points from the experimental data: The number of lipoplexes that attach to the cell surface (4–8), the dose-response curve (transfection efficiency vs. dose), the mean and variance of the onset time of GFP expression are as reported in [44], and the mean maximum GFP expression level. This gives us a good estimate of these five parameters.

The remaining parameters that need optimization are thus the rates of lysis and unpacking. Currently, we don't have a way to distinguish between delays caused by lysis vs. unpacking, so we set unpacking to be immediate. In addition, we assume no negligible degradation of lipoplexes, so we set that rate to a small number.

The values of all parameters, both fixed and fitted, are documented in Table 1. Due to the significant simplifications involved in the model, and the inherent "sloppiness" of models with this many parameters, we do not consider the parameters to be accurate measurements of the real values. The value of the model is demonstrated more by its overall performance and matching with the experimental data.

Model Implementations

The formulation of the SBML implementation of the model is based on reactions, and is a very straightforward step from the reactions documented here. The only difference is the fact that some species are listed, such as End1…End10, instead of the subscripted notation End_i $i = 1…10$ used in the documentation.

The Pi Calculus implementation is discussed in the File S1. This implementation of the model, which was run in SPiM, deals with the variable lipoplex size by including the size as a parameter in the process. It is an implementation of the model in Pi Calculus where the number of lipoplex sizes (the width of the lipoplex size distribution) is restricted to 11, even though 175 is required. This model was run and produced the same data as the Copasi model with only 1 lipoplex size.

The BNGL implementation is discussed in the File S1. This is a prototype of an implementation of the model written in BNGL and run in NFSim. This implementation does not cover enough of the model to produce useful data.

Results and Discussion

Time Courses

Since we are dealing with low copy numbers in the first parts of the transfection process, we need to account for the stochastic nature of them, and see how that compares with a more traditional solution to the equations. Figure 3 shows time courses created by deterministic simulation, i.e. by numerical solution of the differential equations in the green dotted line, and a typical example of time courses created by stochastic simulation, i.e. using Monte Carlo simulation via the Gillespie algorithm in the red full line. The important message in this figure is the very significant difference between deterministic and stochastic simulations. Due to the low copy numbers involved (except for GFP), the deterministic plots are not good representations of the biological reality, and they do not necessarily represent the average behavior of the stochastic simulations. However, they are sometimes useful for running early steps in the parameter estimation task. Figure 3A shows the number of lipoplexes attached to the cell surface, which grows rapidly until the cells are washed after 1 hour of incubation, and then decays exponentially as they enter the cell. Figure 3B shows the number of lipoplexes in endosomes, which demonstrates how they enter and leave the endosomes. Figure 3C shows the number of mRNA molecules, where our example of a stochastic simulation shows that 1 lipoplex (containing 350 mRNA molecules) has entered the cell; this can vary from 0 to about 5. Figure 3D shows the number of GFP molecules, which first increases after mRNA molecules appear and begin to translate, then decreases due to degradation of both mRNA and GFP.

Now that we have set our focus on stochastic simulation time courses, we would like to see how they compare with the experimental data. Figure 4 is another visualization of the GFP time course presented earlier. Figure 4A shows the simulation data. The clustering of the absolute height of the curves results from the fact that mRNA molecules are delivered in "packets", i.e. lipoplexes of size 350. We consider this to be a result of the simplification where we assumed all lipoplexes to contain exactly 350 mRNA molecules, even though the range (within one standard deviation) goes from 270 to 445. This clustering behavior was not observed in the experimental data. The horizontal axis clearly shows the variation in the onset time, and the vertical axis shows the variation in expression level (maximum GFP concentration). These two distributions will be examined in more detail below. Figure 4B shows the experimental data. In the plots, it appears as though the absolute level of GFP expression differs by a factor of 4. However, the value used for parameter optimization was the mean of the maximum GFP expression level, and that is 7.1×10^5 in the experiment and 5.4×10^5 in the simulation. The other values used for optimization varied much less (see Table 1). The time for reaching a peak value in Figure 4B is not easy to see, so we calculated the mean and variance of both distributions, and found that both peak at about 20 hours with a standard deviation of about 5.5 hours.

Simulation vs. Experiment

In order to compare simulation with experiment, probability distributions of some of the key parameters are shown in Figures 5 and 6. In all cases, the experimental data refers to the data published in [44]. Figure 5 shows the onset time of GFP expression, which is defined as the first time where GFP can be detected, and we have measured it by fitting the analytic solution

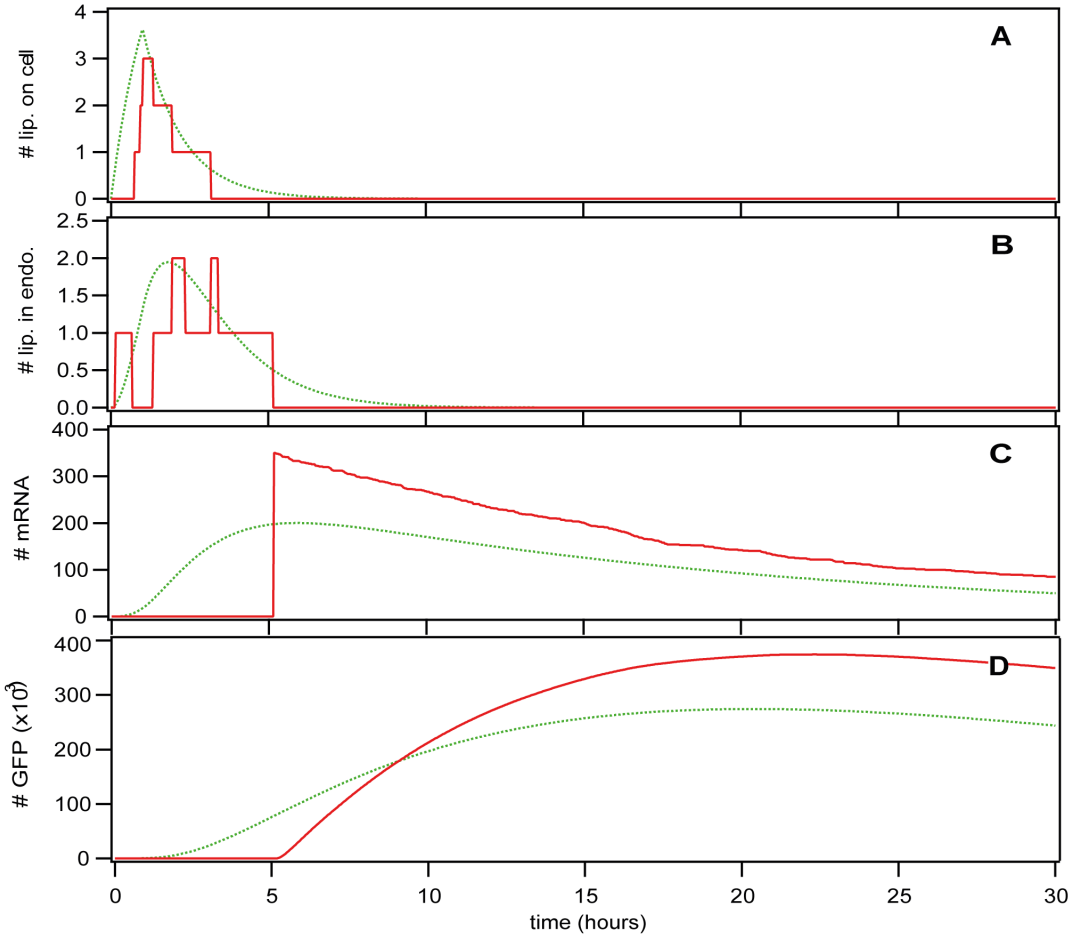

Figure 3. Simulation Time Courses. Green dotted (red full) line: deterministic (stochastic) simulation. A) Number of lipoplexes attached to the cell surface. B) Number of lipoplexes contained in endosomes. C) Number of mRNA molecules in the cell. D) Number of GFP molecules in the cell.
doi:10.1371/journal.pone.0107148.g003

of the expression kinetics to the experimental data and the simulations using the same technique as in the original paper [44]. This makes it unnecessary to use an arbitrary threshold for GFP or to use the simple slope of the curve to determine onset time. The maturation reaction was not included in the original analysis in [44], which means that the maturation delay was included in the onset time there. The green dashed line kM = 9.23 (fitted parameters 3.5 mean and 2.1 width), from literature [48], and solid red line kM = 5.5 (fitted parameters 3.2 mean and 1.6 width), as determined by our parameter optimization. The dotted blue lines show the onset times of the experimental data (fitted with 3.1 mean and 1.5 width). The reason for the difference lies in the fact that all reactions have a small copy number, and thus a large stochastic variation, except for the maturation reaction. We know that, for Poisson processes, the mean is proportional to the number of reactants, and the width is proportional to the square root of the number of reactants, and this number is on the order of 1–100 for endocytosis, 1–100 for lysis, 1–100 for unpacking, 300–2000 for

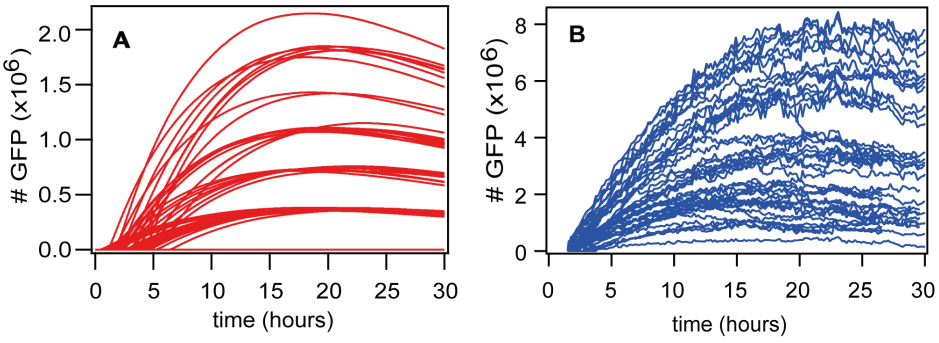

Figure 4. GFP expression: simulation vs. experiment. A) Computer simulation. B) Experimental time courses.
doi:10.1371/journal.pone.0107148.g004

translation, and 200,000–5,000,000 for maturation. In order to match the experimental results, our optimization routine found a maturation rate of 5.5 h^{-1} or 11 min delay. In contrast, the rate of kM = 9.23 (6.5 min) from literature produces a distribution that is too wide. Maturation delays of 20 or 30 minutes also match the experimental data well. This is within the range of published EGFP maturation rates, which vary widely and go as high as a few hours due to the time required for oxidation (more details in File S1). This figure was created in the multiple-lipoplex model, but the streamlined model shows exactly the same behavior, i.e. it is capable of reproducing the experimentally-measured onset time distribution, but also needs the maturation reaction to do so.

Now that we have seen the comparison of simulation and experiment for the onset time of GFP expression, we need to look at how much GFP is created in the cells. Figure 6A shows the distribution of the maximum number of GFP molecules, as determined by fitting the analytical solution of gene expression (translation and degradation) to the data of simulation and experiment. This is the value that we use to determine the level of expression, and, along with the degradation rates, it uniquely determines the time course of GFP expression. The dashed green lines are from a simulation of the streamlined model (fitted with $4.3*10^5$ mean and 0.47 width). The solid red lines are from a simulation of the multiple-lipoplex model (fitted with $5.3*10^5$ mean and 0.69 width). The dotted blue lines show the experimental data (fitted with $7.1*10^5$ mean and 1.1 width). We can see that the simulation of the streamlined model misses the experimental results significantly, which we attribute to the fact that the streamlined model never transports more than one lipoplex per endosome. In contrast to the streamlined model, the multiple-lipoplex model allows a better match to the expression level data. The use of lognormal curves to fit the simulation and experimental data in Figure 6A is more than a convenient guide for the eye; they provide a good representation of the data, since the GFP expression is the result of multiple random processes.

Along with the maximum amount of GFP expressed, we are also interested in seeing how the amount of GFP compares with the dosage of lipoplexes, i.e. the concentration presented to the cells. Figure 6B shows the dose-response relationship, defined as transfection efficiency, i.e. percentage of cells that successfully express GFP vs. concentration of mRNA. The green open triangles are from the simulation of the streamlined model, and the dashed green line is a single-Poissonian fit (fitted parameter 1.1). The open red circles are from the simulation of the multiple-lipoplex model and the solid red line is a double-Poissonian fit (fitted parameters 1.9 and 0.6). The solid blue squares are from the experimental data and the dotted blue line is a double-Poissonian fit (fitted parameters 1.1 and 0.9). In Figure 6B, we can see that the simulation of the streamlined model is much too straight and significantly misses the shape of the experimental results, which we attribute to the fact that the streamlined model never transports more than one lipoplex per endosome. In fact, the good fit of a single Poissonian to the streamlined model is a clear indication that one of the Poissonian processes, representing the number of lipoplexes per endosome, is missing in this model. This process is referred to as L_{eff} in the original paper, and the process that is included in the streamlined model is referred to as N_{eff} [44], File S1. The dose-response relationship for the multiple-lipoplex model shows a reasonable fit to a double Poissonian and to the experimental data, and is a big improvement over the streamlined model.

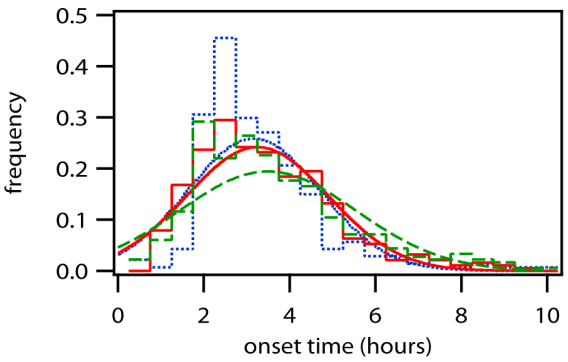

Figure 5. Onset time of GFP expression (Simulation vs. Experiment based on time courses shown in Figure 4). The curves are Gaussian curves based on mean and variance of the full distribution data (shown as a histogram). The dashed green lines show the onset times for simulation with a maturation rate (k_M) of 9.23 taken from literature. The solid red lines show the onset times for simulation with a maturation rate (k_M) of 5.5. The dotted blue lines show the onset times of the experimental data.
doi:10.1371/journal.pone.0107148.g005

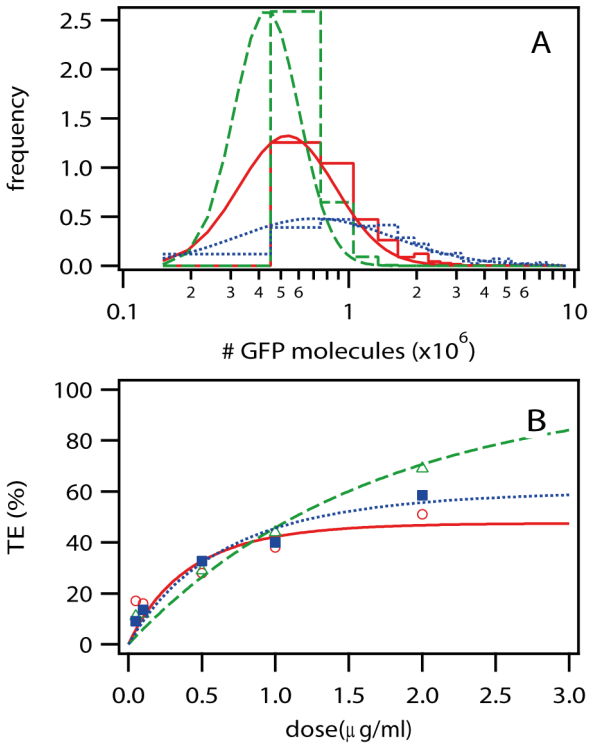

Figure 6. GFP expression (Simulation vs. Experiment based on time courses shown in Figure 4). A) Expression Level. Maximum number of GFP molecules with histograms of the distributions and lognormal fits of the histograms as curves. The dashed green lines are from a simulation of the streamlined model. The solid red lines are from a simulation of the multiple-lipoplex model. The dotted blue lines show the experimental data. B) Dose-Response Relationship. Transfection efficiency (TE) is the percentage of cells that exhibited a successful transfection, based on GFP expression. The curve was determined by varying the dosage (μg/ml) in the experiment, and the initial concentration of lipoplexes in the simulation (L_{ex}). The green open triangles are from the simulation of the streamlined model, and the dashed green line is a single-Poissonian fit. The open red circles are from the simulation of the multiple-lipoplex model and the solid red line is a double-Poissonian fit. The solid blue squares are from the experimental data and the dotted blue line is a double-Poissonian fit.
doi:10.1371/journal.pone.0107148.g006

We can summarize these differences by observing that the streamlined model is capable of reproducing the delay and variance of the onset time of GFP expression, but the multiple-lipoplex model is required to reproduce the dispersion of the data. In other words, multi-level modeling is necessary for reproducing the dispersion of the data, because it is the only model that includes the second Poisson process discussed in the experimental paper.

Predictive Modeling

The power of mathematical modeling is its capability to predict the behavior of systems before running experiments. It is instructive to test the outcome of our simulation for various scenarios of practical relevance in our lab work. In the following, the red circles show the transfection efficiency (percentage of cells transfected) and the green triangles show the maximum GFP expression level.

For determining the dosage presented to the cells, the incubation time, i.e. the time until the cells are washed, plays an important role. Figure 7A shows the transfection efficiency (TE) and the maximal number of eGFP expressed (GFP) as a function of incubation time. The model predicts a strictly linear relation of incubation time and transfection efficiency. This outcome is due to the fact that the model assumes a constant concentration of lipoplexes in bulk and hence a constant diffusion-limited flux. Yet we expect this dependence to be only observable in a very limited time window avoiding saturation of the uptake capacity of the cells as well as the depletion of the lipoplex pool. Most importantly, however, the model does not account for increasing toxic side effects that come with increasing dose.

In this model, the endosome degradation rate is a catch-all for any kind of degradation that occurs before endosomal lysis, especially mRNA degradation, so a small endosome degradation rate should show the benefit of improved mRNA stability. Figure 7B shows the transfection efficiency (TE) and the maximal number of eGFP expressed (GFP) as a function of endosome degradation rate. The solid red and green lines are exponential fits. The exponential increase of transfection efficiency with decreasing degradation rate clearly shows the (expected) benefit of increasing the stability of mRNA. It is interesting to note that the averaged eGFP per expressing cell exhibits a steeper dependence than the fraction of transfected cells (transfection efficiency). When we extrapolate the exponential fits to the point where the endosome degradation rate is zero, we can see that the model predicts approximately 100% transfection efficiency and 1,000,000 maximum GFP for the case of perfectly stable mRNA. Extrapolation to an infinite degradation rate (absolutely unstable mRNA) predicts approximately 0% transfection efficiency as expected. However, this is only approximately 0%, and maximum GFP expression is only calculated for successfully transfected cells, so when we extrapolate to an infinite degradation rate, we see 500,000 GFP molecules per cell, but this is an artifact of the analysis. We should also recall that our model was optimized to an average of 6 lipoplexes adhering to each cell.

In order for the lipoplexes to reach the cytosol and be expressed, they first need to escape from the endosomes, which we have modeled in the endosomal lysis rate. Figure 7C shows the transfection efficiency (TE) and the maximal number of eGFP expressed (GFP) as a function of the lysis rate. The solid red line is an exponential fit while and the solid green line is a linear fit. The increase of transfection efficiency with increasing lysis rate demonstrates the (expected) improvement of transfection with increasing lysis, or endosomal escape [4,9,49–52]. We expect a similar effect when changing the attach rate via the use of magnetofection [8].

The size of the lipoplexes may have an important influence on their uptake. Figure 7D shows the transfection efficiency (TE) and the maximal number of eGFP expressed (GFP) as a function of the lipoplex size. We can see that the model predicts a higher percentage of cells transfected when the lipoplexes are smaller (but total mRNA concentration kept the same), and a higher total amount of GFP when the lipoplexes are larger. This opposing

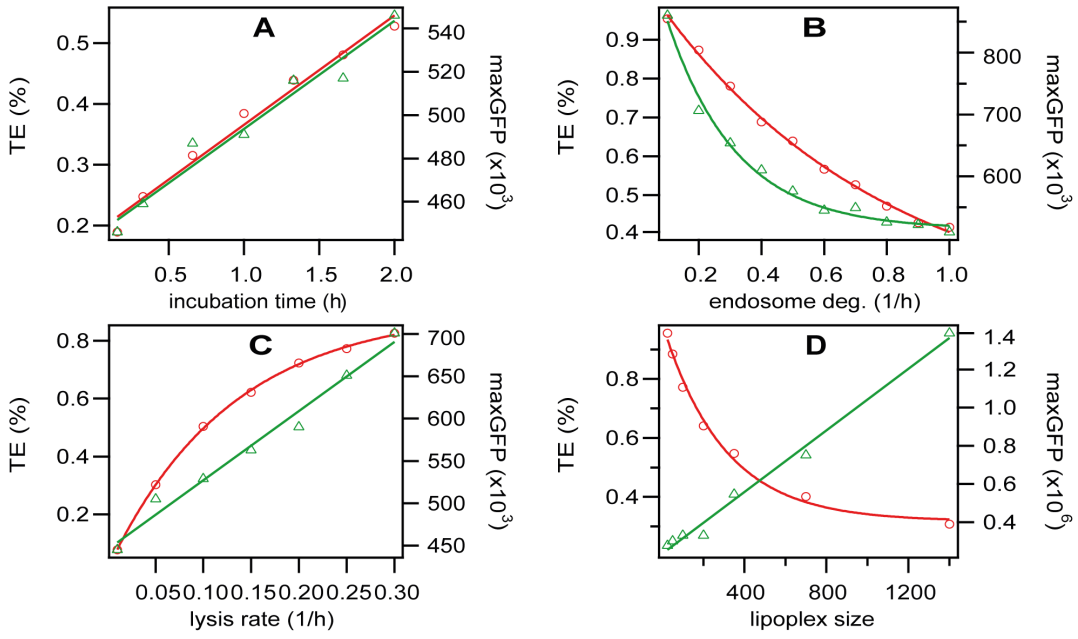

Figure 7. Predictive Modeling. All plots show a parameter vs. transfection efficiency (TE, red circles) and protein expression (GFP, green triangles). The lines are linear or exponential fits. A) Incubation time. B) Endosome degradation rate. C) Lysis rate. D) Lipoplex size.
doi:10.1371/journal.pone.0107148.g007

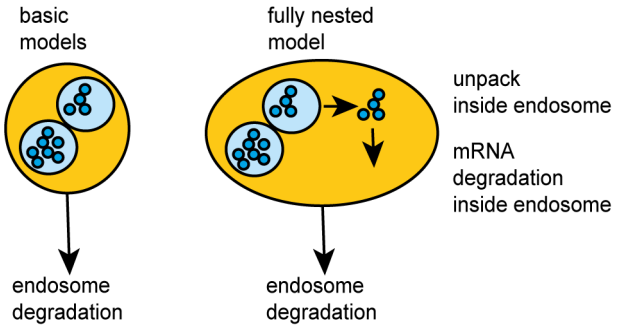

Figure 8. Key aspect of the fully nested transfection model. In addition to the processes in the multiple-lipoplex model, the fully nested model includes unpacking of lipoplexes and degradation of mRNA within endosomes.
doi:10.1371/journal.pone.0107148.g008

effect occurs because we assume a constant uptake rate independent of size and smaller lipoplexes mean a larger number of them, which increases the probability of successful transfection, while larger lipoplexes are capable of transporting more material. A size-independent uptake rate, however, is taken with a very big caveat. In fact, the dependence of uptake on size has been shown in experiment for gold nanoparticles [53–56]. Yet, there is some value to the finding that in case of variation of experiments focused on an optimal lipoplex size, in which case the size dependence might be weak, transfection efficiency and GFP expression react in the opposite direction.

Conclusions and Outlook

We have presented a kinetic model for mRNA delivery via transfection of lipoplexes. The model consists of a chain of transfer events including lipoplex attachment, endocytosis, endosomal lysis, unpacking, translation and maturation. It was shown that parameter estimation allows direct comparison to the outcome of a single-cell transfection analysis. The model provides a kinetic model that reproduces both the delay and dispersion of the onset time and also the dose-response relationship. The delay can be reproduced using the streamlined model, but the multiple-lipoplex model, which is based on multi-level modeling, is necessary in order to reproduce the dispersion of the data. The key findings are that in order to achieve the observed level of GFP expression, as expressed in the maxGFP distribution, we need to use the multiple-lipoplex model. A multiple-lipoplex model achieves the correct width (stochastic variance) of the probability distribution for the onset time of GFP expression if the maturation reaction is included. A hallmark of the multiple-lipoplex model is its combinatorial manifold, which exceeds the capacity of ordinary modeling platforms. We showed that a reduction of the combinatorial space to a limited variance was able to approximate the shape of the dose-response relationship.

Extensions of the model that might be necessary as more refined data become available are more explicit rate equations that include cooperative behavior (Hill kinetics) or e.g. enzyme limited reactions (Michaelis Menten type kinetics). Furthermore, degradation processes could be broken down into specifically known pathways. Yet the most important uncertainty concerns the uptake process itself. The fact that we used a single, uniform rate of attachment of lipoplexes to clathrin-coated pits and that the rate of endocytosis in our model does not depend on the size of the pit is first of all due to missing quantitative data. We have assumed that endosomes first undergo lysis, then the lipoplexes are unpacked, and then the mRNA can begin translation and degradation. However, unpacking might occur within the endosome before lysis and, as mentioned earlier, mRNA degradation might begin in the endosome before lysis. Furthermore, we don't currently have a way to distinguish between a delay caused by lysis and delay caused by unpacking, so we have simplified the model to treat unpacking as an immediate reaction.

A key aspect of this investigation is multi-level modeling, which leads to a combinatorial explosion of variables and reactions, but this could be solved more elegantly by a computational system that copes with it directly. However, this does not make the combinatorial explosion disappear; the burden is simply transferred from the user to the tool in the form of dynamic creation of species. The basis for this already exists in SBML, Copasi, SPiM, BioNetGen, NFsim, and ML-Rules, which introduces the concept of nested species [57,58], meaning that one species, such as mRNA molecules, can exist and exhibit behavior within another species, such as a lipoplex or endosome. This would make it possible to formulate the model in a more elegant way, which would be easier to understand. As a second benefit, it would make it possible to remove a significant limitation of today's model, which assumes that all lipoplexes have the same size and leads to a clustering of GFP expression levels visible in Figure 4, and it would be possible to model explicit unpacking of lipoplexes and degradation of mRNA within endosomes, instead of resorting to an endosome degradation reaction, as shown in the fully nested model (Figure 8). Finally, it would also make it possible to use species as building blocks to create new ones; for example, chemical reaction networks could be used to build organelles, which could be used to build cells, etc. This type of model is often required for nanoparticle transport in general, and should provide a basis for more predictive modeling in that area.

Beside all well-founded shortcomings of the current model limitations, there is substantial value added by comparison of modeling and experimental data. The fact that data are reproduced by a set of parameters that is optimized by the same number of experimental determinants justifies our assertion that the model has significant predictive power. We have done predictive modeling by analyzing the effect of varying parameters, and the results either agree with existing experimental data (e.g. dose-response), confirm known aspects (e.g. importance of endosomal escape), or predict new effects, such as the effect that decreasing the size of the lipoplexes has on transfection efficiency and GFP expression.

With appropriate modifications, this model should be useful for new experimental work. The key parameters include the rates of attachment, endocytosis, lysis, unpacking, and the size-dependency of those rates; as new data on these parameters becomes available, this should lead to a significant improvement in the quality of the model.

Supporting Information

File S1
Code S1. Script for automated simulation of dose-response relationship.
Code S2. script for automated simulation of lipoplex size dependency.
Code S3. C# source code for program to set parameters in Copasi model.
Code S4. C# source code for program to run Copasi model multiple times and analyze results in Igor Pro.
Code S5. C# source code for program to run TFC.exe and optimize via simulated annealing algorithm.

Code S6. C# source code for program to run Copasi streamlined (reduced) model multiple times and analyze results in Igor Pro.
Code S7. C# source code for program to run TFRC.exe and optimize via simulated annealing algorithm.
Code S8. Igor Pro procedure for analyzing results of Copasi model (TFC.cps and TFC.cps).
Code S9. Perl script for running SPiM model.
Code S10. Igor Pro procedure for analyzing results of SPiM model.
Code S11. Igor Pro procedure for creating figures.
Dataset S1. Dose-response data (Figure 6).
Dataset S2. GFP data (Figure 4).
Dataset S3. Lipoplex size data (Figure 7).
Dataset S4. Max GFP experiment (Figure 5B).
Dataset S5Max GFP reduced model (Figure 5B).
Dataset S6. Max GFP (Figure 5B).
Dataset S7. Onset time experiment (Figure 5A).
Dataset S8. Onset time reduced model (Figure 5A).
Dataset S9. Onset time (Figure 5A).
Dataset S10. Time courses (Figure 3).
Model S1. Copasi model for deterministic simulation of multiple lipoplex model.
Model S2. SBML model for deterministic simulation of multiple lipoplex model.
Model S3. Copasi model for stochastic simulation of multiple lipoplex model.
Model S4. SBML model for stochastic simulation of multiple lipoplex model.
Model S5. Copasi model for deterministic simulation of streamlined (reduced) model.
Model S6. SBML model for deterministic simulation of streamlined (reduced) model.
Model S7. Copasi model for stochastic simulation of streamlined (reduced) model.
Model S8. SBML model for stochastic simulation of streamlined (reduced) model.
Model S9. SPiM model.
Model S10. Version 1 of BNGL (BioNetGenML) model for NFSim.
Model S11. SBML model for BNGL.
Text S1. Detailed model description.
Text S2. ODEs as created by Copasi in TeX format.
Text S3. ODEs imbedded in LaTeX document file.
Text S4. ODEs in PDF format (from LaTeX).
(ZIP)

Author Contributions

Conceived and designed the experiments: TL CL. Performed the experiments: TL. Analyzed the data: TL. Contributed reagents/materials/analysis tools: TL. Wrote the paper: TL CL JR.

References

1. Neumann B, Held M, Liebel U, Erfle H, Rogers P, et al. (2006) High-throughput RNAi screening by time-lapse imaging of live human cells. Nat Meth 3: 385–390.
2. Rantala JK, Mäkelä R, Aaltola A-R, Laasola P, Mpindi J-P, et al. (2011) A cell spot microarray method for production of high density siRNA transfection microarrays. BMC genomics 12: 162.
3. Rinaudo K, Bleris L, Maddamsetti R, Subramanian S, Weiss R, et al. (2007) A universal RNAi-based logic evaluator that operates in mammalian cells. Nat Biotech 25: 795–801.
4. Tseng Y-C, Mozumdar S, Huang L (2009) Lipid-based systemic delivery of siRNA. Advanced drug delivery reviews 61: 721–731.
5. Debus H, Baumhof P, Probst J, Kissel T (2010) Delivery of messenger RNA using poly(ethylene imine)–poly(ethylene glycol)-copolymer blends for polyplex formation: Biophysical characterization and in vitro transfection properties. Journal of Controlled Release 148: 334–343.
6. Kamiya H, Akita H, Harashima H (2003) Pharmacokinetic and pharmacodynamic considerations in gene therapy. Drug Discovery Today 8: 990–996.
7. Malone RW, Felgner PL, Verma IM (1989) Cationic liposome-mediated RNA transfection. Proceedings of the National Academy of Sciences 86: 6077–6081.
8. Sauer AM, de Bruin KG, Ruthardt N, Mykhaylyk O, Plank C, et al. (2009) Dynamics of magnetic lipoplexes studied by single particle tracking in living cells. Journal of Controlled Release 137: 136–145.
9. Schloßbauer A, Sauer AM, Cauda V, Schmidt A, Engelke H, et al. (2012) Cascaded Photoinduced Drug Delivery to Cells from Multifunctional Core–Shell Mesoporous Silica. Advanced Healthcare Materials 1: 316–320.
10. Tachibana R, Harashima H, Shinohara Y, Kiwada H (2001) Quantitative studies on the nuclear transport of plasmid DNA and gene expression employing nonviral vectors. Advanced drug delivery reviews 52: 219–226.
11. Mac Gabhann F, Annex BH, Popel AS (2010) Gene therapy from the perspective of systems biology. Current opinion in molecular therapeutics 12: 570.
12. Morille M, Passirani C, Vonarbourg A, Clavreul A, Benoit J-P (2008) Progress in developing cationic vectors for non-viral systemic gene therapy against cancer. Biomaterials 29: 3477–3496.
13. Nguyen J, Szoka FC (2012) Nucleic Acid Delivery: The Missing Pieces of the Puzzle? Accounts of Chemical Research 45: 1153–1162.
14. Pedroso de Lima MC, Simões S, Pires P, Faneca H, Düzgüneş N (2001) Cationic lipid–DNA complexes in gene delivery: from biophysics to biological applications. Advanced drug delivery reviews 47: 277–294.
15. Rädler JO, Koltover I, Salditt T, Safinya CR (1997) Structure of DNA-Cationic Liposome Complexes: DNA Intercalation in Multilamellar Membranes in Distinct Interhelical Packing Regimes. Science 275: 810–814.
16. Safinya CR (2001) Structures of lipid–DNA complexes: supramolecular assembly and gene delivery. Current Opinion in Structural Biology 11: 440–448.
17. Schaffert D, Wagner E (2008) Gene therapy progress and prospects: synthetic polymer-based systems. Gene Ther 15: 1131–1138.
18. Shahrezaei V, Swain PS (2008) Analytical distributions for stochastic gene expression. Proceedings of the National Academy of Sciences 105: 17256–17261.
19. Varga CM, Hong K, Lauffenburger DA (2001) Quantitative Analysis of Synthetic Gene Delivery Vector Design Properties. Mol Ther 4: 438–446.
20. Varga CM, Tedford NC, Thomas M, Klibanov AM, Griffith LG, et al. (2005) Quantitative comparison of polyethylenimine formulations and adenoviral vectors in terms of intracellular gene delivery processes. Gene Ther 12: 1023–1032.
21. Dinh AT, Pangarkar C, Theofanous T, Mitragotri S (2007) Understanding intracellular transport processes pertinent to synthetic gene delivery via stochastic simulations and sensitivity analyses. Biophys J 92: 831–846.
22. Summers HD, Rees P, Holton MD, Rowan Brown M, Chappell SC, et al. (2011) Statistical analysis of nanoparticle dosing in a dynamic cellular system. Nat Nano 6: 170–174.
23. Schwake G, Youssef S, Kuhr JT, Gude S, David MP, et al. (2010) Predictive modeling of non-viral gene transfer. Biotechnol Bioeng 105: 805–813.
24. Locke JC, Elowitz MB (2009) Using movies to analyse gene circuit dynamics in single cells. Nat Rev Microbiol 7: 383–392.
25. Fisher J, Henzinger TA (2007) Executable cell biology. Nat Biotechnol 25: 1239–1249.
26. Mahmutovic A, Fange D, Berg OG, Elf J (2012) Lost in presumption: stochastic reactions in spatial models. Nature Methods 9: 1163–1166.
27. Hucka M, Bergmann F, Hoops S, Keating S, Sahle S, et al. (2010) The Systems Biology Markup Language (SBML): Language Specification for Level 3 Version 1 Core. Nature Precedings.
28. Funahashi A, Morohashi M, Kitano H, Tanimura N (2003) CellDesigner: a process diagram editor for gene-regulatory and biochemical networks. Biosilico 1: 159–162.
29. Hoops S, Sahle S, Gauges R, Lee C, Pahle J, et al. (2006) COPASI–a COmplex PAthway SImulator. Bioinformatics 22: 3067–3074.
30. Milner R (1999) Communicating and Mobile Systems: The Pi Calculus: Cambridge University Press.
31. Priami C (1995) Stochastic π-calculus. The Computer Journal 38: 578–589.
32. Priami C, Regev A, Shapiro E, Silverman W (2001) Application of a stochastic name-passing calculus to representation and simulation of molecular processes. Information processing letters 80: 25–31.
33. Regev A, Shapiro E (2002) Cellular abstractions: Cells as computation. Nature 419: 343–343.
34. Regev A, Shapiro E (2004) The π-calculus as an abstraction for biomolecular systems. Modelling in Molecular Biology: 219–266.
35. Regev A, Silverman W, Shapiro E (2001) Representation and simulation of biochemical processes using the-calculus process algebra. pp. 459–470.
36. Regev A, Panina EM, Silverman W, Cardelli L, Shapiro E (2004) BioAmbients: an abstraction for biological compartments. Theoretical Computer Science 325: 141–167.

37. Guerriero M, Priami C, Romanel A (2007) Modeling static biological compartments with beta-binders. Algebraic Biology: 247–261.
38. Guerriero ML, Priami C, Romanel A (2006) Beta-binders with static compartments.
39. Priami C, Quaglia P (2005) Beta binders for biological interactions. Springer.pp. 20–33.
40. Phillips A, Cardelli L. Efficient, correct simulation of biological processes in the stochastic pi-calculus; 2007. Springer.pp. 184–199.
41. Faeder JR, Blinov ML, Hlavacek WS (2009) Rule-based modeling of biochemical systems with BioNetGen. Systems biology: Springer.pp. 113–167.
42. Sneddon MW, Faeder JR, Emonet T (2011) Efficient modeling, simulation and coarse-graining of biological complexity with NFsim. Nat Meth 8: 177–183.
43. Dobay MPD, Alberola AP, Mendoza ER, Rädler JO (2012) Modeling nanoparticle uptake and intracellular distribution using stochastic process algebras. Journal of Nanoparticle Research 14.
44. Leonhardt C, Schwake G, Stögbauer TR, Rappl S, Kuhr J-T, et al. (2013) Single-cell mRNA transfection studies: delivery, kinetics and statistics by numbers. Nanomedicine: Nanotechnology, Biology and Medicine.
45. Andries O, De Filette M, Rejman J, De Smedt SC, Demeester J, et al. (2012) Comparison of the Gene Transfer Efficiency of mRNA/GL67 and pDNA/GL67 Complexes in Respiratory Cells. Molecular Pharmaceutics 9: 2136–2145.
46. Tavernier G, Andries O, Demeester J, Sanders NN, De Smedt SC, et al. (2011) mRNA as gene therapeutic: How to control protein expression. Journal of Controlled Release 150: 238–247.
47. Khalil IA, Kogure K, Akita H, Harashima H (2006) Uptake pathways and subsequent intracellular trafficking in nonviral gene delivery. Pharmacological reviews 58: 32–45.
48. Megerle JA, Fritz G, Gerland U, Jung K, Rädler JO (2008) Timing and Dynamics of Single Cell Gene Expression in the Arabinose Utilization System. Biophys J 95: 2103–2115.
49. Chan C-L, Majzoub RN, Shirazi RS, Ewert KK, Chen Y-J, et al. (2012) Endosomal escape and transfection efficiency of PEGylated cationic liposome–DNA complexes prepared with an acid-labile PEG-lipid. Biomaterials 33: 4928–4935.
50. Dominska M, Dykxhoorn DM (2010) Breaking down the barriers: siRNA delivery and endosome escape. Journal of cell science 123: 1183–1189.
51. Sauer AM, Schlossbauer A, Ruthardt N, Cauda V, Bein T, et al. (2010) Role of endosomal escape for disulfide-based drug delivery from colloidal mesoporous silica evaluated by live-cell imaging. Nano Lett 10: 3684–3691.
52. Xu Y, Szoka FC (1996) Mechanism of DNA Release from Cationic Liposome/DNA Complexes Used in Cell Transfection†,‡. Biochemistry 35: 5616–5623.
53. Chithrani BD, Chan WCW (2007) Elucidating the Mechanism of Cellular Uptake and Removal of Protein-Coated Gold Nanoparticles of Different Sizes and Shapes. Nano Lett 7: 1542–1550.
54. Chithrani BD, Ghazani AA, Chan WCW (2006) Determining the Size and Shape Dependence of Gold Nanoparticle Uptake into Mammalian Cells. Nano Lett 6: 662–668.
55. Gao H, Shi W, Freund LB (2005) Mechanics of receptor-mediated endocytosis. Proceedings of the National Academy of Sciences of the United States of America 102: 9469–9474.
56. Jiang W, KimBetty YS, Rutka JT, ChanWarren CW (2008) Nanoparticle-mediated cellular response is size-dependent. Nat Nano 3: 145–150.
57. Faeder JR (2011) Toward a comprehensive language for biological systems. BMC biology 9: 68.
58. Maus C, Rybacki S, Uhrmacher AM (2011) Rule-based multi-level modeling of cell biological systems. BMC Syst Biol 5: 166.
59. Zhou J, Yockman J, Kim S, Kern S (2007) Intracellular Kinetics of Non-Viral Gene Delivery Using Polyethylenimine Carriers. Pharmaceutical Research 24: 1079–1087.
60. Kremers G-J, Goedhart J, van den Heuvel DJ, Gerritsen HC, Gadella TWJ (2007) Improved Green and Blue Fluorescent Proteins for Expression in Bacteria and Mammalian Cells. Biochemistry 46: 3775–3783.
61. Reid BG, Flynn GC (1997) Chromophore Formation in Green Fluorescent Protein. Biochemistry 36: 6786–6791.
62. Sniegowski JA, Lappe JW, Patel HN, Huffman HA, Wachter RM (2005) Base catalysis of chromophore formation in Arg96 and Glu222 variants of green fluorescent protein. Journal of Biological Chemistry 280: 26248–26255.
63. Tsien RY (1998) The green fluorescent protein. Annu Rev Biochem 67: 509–544.

6 Outlook

In this thesis, the dynamics of mRNA-mediated protein expression was quantitatively investigated at the single-cell level and a new model for mRNA transfection was introduced. This way, the distributions of the translational efficiency and the life times of mRNA as well as EGFP were determined in single cells. Additionally, we were able to predict transfection efficiencies and protein levels for varying rates that are characteristic for the mRNA-transfer process on the basis of the new transfection model. Combining single-cell arrays with automated image analysis tools has great potential for high-throughput, quantitative transfection studies in general. More specifically, this approach could also be used for systematic studies of siRNA knockdown that are carried out in our laboratory.

For the further development of predictive models, gene carrier systems that offer predictable transfer kinetics are desirable. We have tested various alternative liposome formulations in order to produce our own vectors for transfection. Because we observed severe toxic side effects, we decided to further rely on the standard commercial liposome-based transfection reagent *Lipofectamine 2000*®. Recently, promising advances in the development of photo-inducible gene carriers and lipoplexes with tunable cargo loads down to single RNA strands have been reported [121, 158]. In the future, such vectors could be used. Ideally, a database of cell-type specific rates for intracellular trafficking could be created using such vectors. Tunable gene delivery frameworks could also provide quantitative data for modeling complex gene-regulatory systems and for RNA-based logic circuits similar to those described in [95, 96, 159].

So far, all data fitting was done using *IGOR Pro* software, whose curve fitting tool is based on the Levenberg-Marquardt algorithm. This algorithm finds optimal parameter values by minimising an error function, which is a measure for the distance between measured and fitted values. In our case, *IGOR Pro* used a local optimisation method. Because this estimated optimum is influenced by the initial parameter guess, it is possible that the best fit value for a parameter is a local, however not the global optimum in parameter space. To overcome this problem, a multi-start, global optimisation method like the simulated annealing algorithm used in chapter 5 was tested on existing data sets in collaboration with the institute of computational biology (Helmholtz Zentrum München). We intend to continue on this path and to also test alternative fitting methods that assume parameter distributions or that simultaneously fit multiple single-cell time courses.

Furthermore, the simple biochemical rate model that was employed for this thesis is symmetric in mRNA and EGFP degradation rates and assumes a simple model of exponential mRNA decay. We

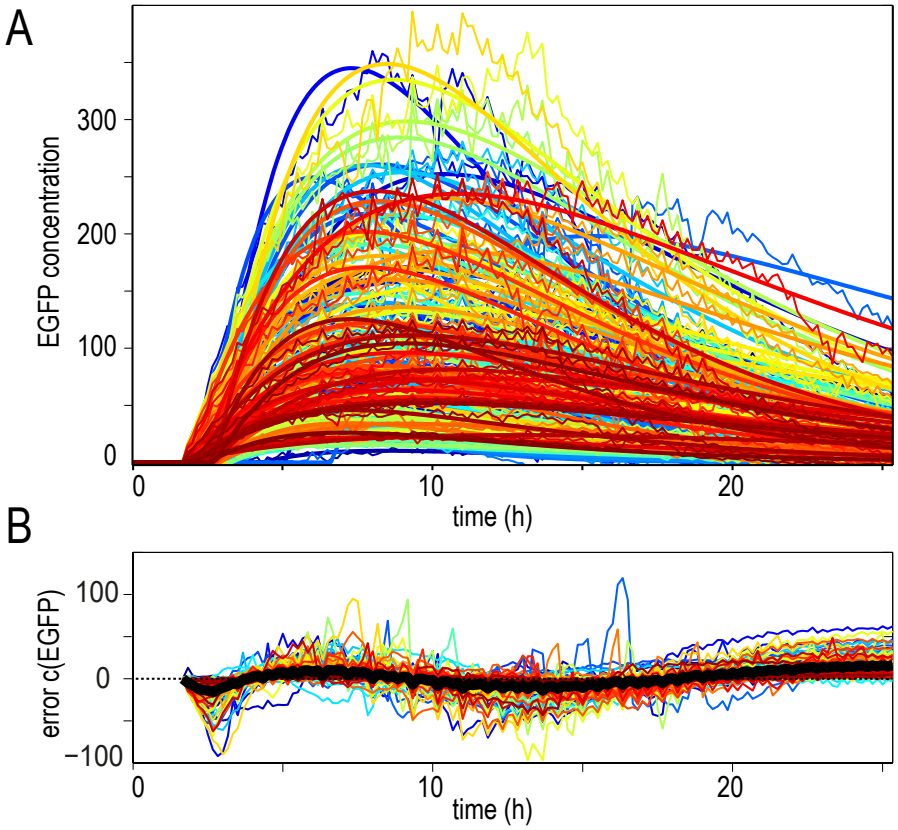

Figure 6.1: *(A) Exemplary fits of single-cell protein expression time courses. (B) The fit residuals show small, but systematic variations with time (the mean residual is shown in black). We intend to eliminate these deviations by more complex modeling approaches and a refined experimental setup. Figure reprinted with permission of F. Fröhlich, Helmholtz Zentrum München.*

have reason to believe that this might not be the complete picture to describe mRNA-mediated protein expression. To counter these problems, we started to test more complex and thus possibly more realistic models. These include scenarios where the degradation rates are no longer interchangeable. Also, a model that includes a limited pool of ribosomes, where ribosomal (un)binding to mRNA is represented by an explicit rate is tested. The overall goal is to find a model that yields the best fit to our data while having the least parameters.

If it were possible to artificially rule out one of the decay channels, it might be possible to identify the contribution of specific degradation channels, which have been described in [67], to the overall mRNA decay. For example, this could be achieved by stabilising the mRNA at one of its ends so that degradation from this side is slowed down drastically as compared to the other decay pathways.

We found systematic variations with time of the residuals for the current model (see Figure 6.1). This is observed both for the fits of single-cell time courses and for the mean fit residuals of many time courses. Also, signal-to-noise-related effects remain a major challenge. External disturbances such as uneven illumination and fluorescence lamp intensity fluctuations can impair data quality and should be reduced to a minimum level. This can be achieved by using illumination systems with feedback-loops or light emitting diode (LED)-based fluorescence excitation. Until recently,

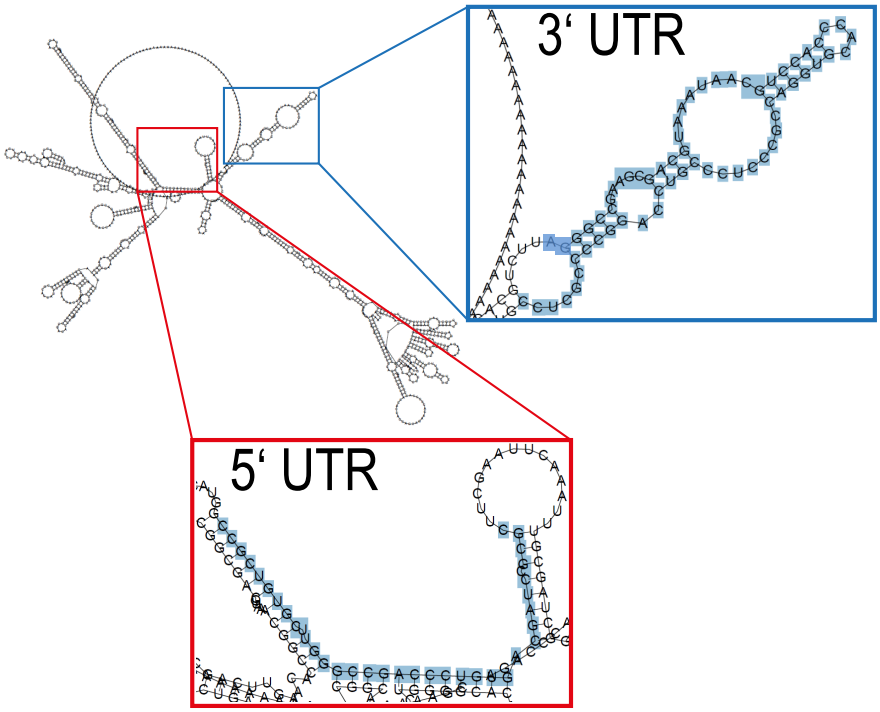

Figure 6.2: *Exemplary secondary structure of a 5’UTR and 3’UTR stabilised mRNA construct encoding for d2EGFP as used for the experiments in chapter 4. Enlarged areas show the UTR sequences.*

LED light sources had the disadvantage that long exposure times, which may be harmful to the cells, were needed because of their low intensities. Thanks to the rapid development in this field, promising systems for live-cell imaging applications are now available. A white light LED-based illumination unit will thus soon be included in our setup. Additionally, the background correction implementation is currently being improved in order to comprehensively take the effects of uneven illumination into account.

To our surprise, we observed positively correlated mRNA and protein life times. This effect might be caused by a cell cycle related phenomenon. If the number of degradation enzymes depends on the cell cycle, this could in turn affect the rates of protein expression, which would be a possible explanation for the observed correlations. To further investigate this, we intend to either use cell cycle markers or to synchronise cells in future experiments. Furthermore, we plan to statistically analyse the effect of different mRNA-cappings. By comparing differently stabilised mRNA constructs, we hope to gain insight into the working mechanisms of capping structures, which would be of interest for the field of systems biology.

Finally, the impact of mRNA-secondary structures on their life times is a topic that is definitely worth investigating. Figure 6.2 shows one of all possible secondary structures of an mRNA construct that was used for the experiments of chapter 4. The mRNA structure was predicted by a freely available software that minimises free energy (*RNA fold*, [160–162]). In the future, features of the mRNA secondary structure, e.g. stem-loops, could be related to experimentally observed alterations of mRNA stability. Also, backfolding of inserted UTRs onto the coding sequence could

be identified with the help of mRNA structure prediction software. This phenomenon might explain differences in the stability ranking of different constructs that have been found for different reporter genes. Exchanging only a few bases in the mRNA sequence can cause significant alterations in the mRNA secondary structure, which might in turn change the stability and hence the life time of the mRNA.

The experimental and the modeling techniques that have been presented in this thesis may help to gain a deeper understanding of mRNA transfer and turnover, thus contributing to the framework of combined quantitative, high-throughput single-cell studies and predictive modeling. Hopefully, they will also help to further develop and characterise artificial vectors for gene delivery and to identify superior mRNA constructs that are relevant for gene therapy.

A Further publication P3: Cell motility in 3D-microstructured environments

The results of the following publication are based on a project that I was involved in during the first year of my doctoral thesis. More specifically, I conducted the experiments on migration in microstructured pillar islands and analysed the corresponding data (see Figure 3 of [163]).
In this project, the influence of mechanical stimuli on the migration modes of highly motile cells was studied. To this aim, cell migration on flat and 3D structured substrates that consisted of micropillar-arrays was investigated. Cell trajectories were analysed with high time-resolution by a two-state motility model. Most importantly, we found that the cells are attracted by the additional surface that is available for them in the vicinity of micropillars.

Further publication P3

Contact-controlled amoeboid motility induces dynamic cell trapping in 3D-microstructured surfaces

By

Delphine Arcizet, Sofia Capito, Mari Gorelashvili, Carolin Leonhardt, Marion Vollmer, Simon Youssef, Susanne Rappl, Doris Heinrich

published in
Soft Matter 2012,8,1473-1481.
doi:10.1039/c1sm05615h

Soft Matter

Cite this: *Soft Matter*, 2012, **8**, 1473

www.rsc.org/softmatter

PAPER

Contact-controlled amoeboid motility induces dynamic cell trapping in 3D-microstructured surfaces†

Delphine Arcizet,‡§ Sofia Capito,‡ Mari Gorelashvili, Carolin Leonhardt, Marion Vollmer, Simon Youssef, Susanne Rappl and Doris Heinrich*

Received 7th April 2011, Accepted 7th November 2011
DOI: 10.1039/c1sm05615h

On flat substrates, several cell types exhibit amoeboid migration, which is characterized by restless stochastic successions of pseudopod protrusions. The orientation and frequency of new membrane protrusions characterize efficient search modes, which can respond to external chemical stimuli as observed during chemotaxis in amoebae. To quantify the influence of mechanical stimuli induced by surface topography on the migration modes of the amoeboid model organism *Dictyostelium discoideum*, we apply high resolution motion analysis in microfabricated pillar arrays of defined density and geometry. Cell motion is analyzed by a two-state motility-model, distinguishing directed cellular runs from phases of isotropic migration that are characterized by randomly oriented cellular protrusions. Cells lacking myosin II or cells deprived of microtubules show significantly different behavior concerning migration velocities and migrational angle distribution, without pronounced attraction to pillars. We conclude that microtubules enhance cellular ability to react with external 3D structures. Our experiments on wild-type cells show that the switching from randomly formed pseudopods to a stabilized leading pseudopod is triggered by contact with surface structures. These alternating processes guide cells according to the available surface in their 3D environment, which we observed dynamically and in steady-state situations. As a consequence, cells perform "home-runs" in low-density pillar arrays, crawling from pillar to pillar, with a characteristic dwell time of ~75 s. At the boundary between a flat surface and a 3D structured substrate, cells preferentially localize in contact with micropillars, due to the additionally available surface in the microstructured arrays. Such responses of cell motility to microstructures might open new possibilities for cell sorting in surface structured arrays.

Introduction

Amoeboid migration is studied preferentially in the social amoeba *Dictyostelium discoideum* or in neutrophils. This migration type is characterized by formation of stochastically generated pseudopods and proves to be extremely efficient in hunting for food or invadors.[1] The fact that external stimuli affect the cellular search behavior is a major feature of this seemingly random cell motion. Membrane proteins sense and integrate external cues and transport signals to downstream effectors for specific cellular response or adaption.[2,3]

A prominent example for adaption to external stimuli is chemotaxis, where cells exposed to a gradient of chemoattractant molecules move up-gradient.[4] During the past decade, many biochemical signalling pathways involved in chemotaxis have been elucidated for the motile cell archetype *Dictyostelium discoideum*.[5–8] The recruitment of key cytosolic proteins to the plasma membrane of the leading pseudopod triggers directional actin polymerization and subsequent migration towards the chemoattractant source. Analysis of cellular migration patterns under external stimuli reveals biased random search strategies with increased migration efficiency in terms of speed and directional persistence towards the chemotactic source.[9]

However, in the absence of chemotactic signaling, *D. discoideum* cells show a "basic motile behavior".[10] The dynamics of this random migration mode is not fully understood yet. Even less advanced is the general understanding of amoeboid migration in the presence of three dimensional surface features.

Cell velocity and shape analysis led to simple models of amoeboid migration, describing it as a random walk.[10] Recently, a refined view of amoeboid locomotion emerged,

Center for Nanoscience (CeNS) and Faculty of Physics, Ludwig-Maximilians-Universität München, Geschwister-Scholl-Platz 1, 80539 München, Germany. E-mail: doris.heinrich@lmu.de

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c1sm05615h

‡ These two authors contributed equally to this work.

§ Current address: Institut de Biologie Structurale Jean-Pierre Ebel, Commissariat à l'Energie Atomique, 41 rue Jules Horowitz, 38027 Grenoble, France.

based on detailed analyses revealing correlations between cellular orientation and migration direction[11–14] as well as motional persistence.[15] The current understanding of basic amoeboid locomotion assumes two alternating motility modes: a random probing mode (*rm*-mode) and a fast, directed running mode (*dir*-run).[16] Both modes originate from the spatiotemporal behavior of biochemical networks in the cell cytoskeleton: spontaneous actin polymerization near the plasma membrane results in random protrusions of multiple pseudopods (polypodial state, *rm*-mode), until one of them is reinforced, suppressing other pseudopods and temporarily stabilizing a directed run (monopodial state or *dir*-run).[17] This two-state-motility in spontaneous amoeboid migration is strikingly similar to motion patterns found in large animals,[18] which exhibit a local search mode and a long-distance running mode, and to the "run and tumble" motion of prokaryotic organisms,[19] which emphasizes the universal character of this very efficient search strategy.

Moreover, *D. discoideum* cells were found to be capable of sensing physical constraints, an obviously vital asset for migration in a natural soil microenvironment.[20] In fact, upon mutation of adhesion proteins,[21] a complex topographical environment causes a drastic decrease in migration efficiency, in ways that go totally unnoticed on flat substrates.[22] Studies on the effect of topographical cues on cytoskeleton dynamics have been facilitated by recent advances in microfabrication. The strong influence of prepatterned surfaces shaped as grooves or pillars on cell polarization and migration has been demonstrated for a variety of cell types.[23–28] Mahmud *et al.* successfully sorted metastatic cells from healthy cells by directed migration in microstructured ratchets.[29] Using antibody-coated microstructures in a fluidic system, rare circulating tumour cells could even be isolated from blood.[30] Thus topographical microstructured setups have not only been demonstrated to constitute useful tools for medical assays, but they also allow the analysis of amoeboid migration in an environment simulating the *in vivo* situation. The differences between cell migration on flat substrates and within a three-dimensional topography are still unknown and remain a topic of ongoing investigations, for individual cells as well as for cell sheets.[31]

In this work, we analyse the migrational behavior of *D. discoideum* on three-dimensional and flat substrates in order to unravel the role of surface structures on motility modes. We find that microstructured surfaces are not sensed as obstacles, preventing cells from progressing efficiently, but rather lead to cell deflection or trapping, depending on the cells' initial motility mode and the number of surface structures in contact with the cells.

Materials and methods

We studied the migration modes of single *D. discoideum* cells in the vegetative state, in the absence of chemoattractant molecules and at low cell densities. Substrates for the migration assays were fabricated from the transparent polymer polydimethylsiloxane (PDMS) and consisted of micropillar arrays in combination with flat surfaces (which were used as a reference to study spontaneous cell motility in a homogeneous 2D environment).

1. Cell culture and observation

We used a mutant *D. discoideum* cell line expressing free GFP, imaged the cells by brightfield and fluorescence microscopy, and recorded time-lapse movies of 200 to 600 frames at a frame rate of 0.1 Hz.

***Dictyostelium discoideum* culture and preparation for experiments.** The axenic *D. discoideum* strain with GFP insertion, strain HG1694,[32] and the myosin II-null mutant with LimEΔcc-GFP, strain HS2205,[33] were obtained from Dr Günther Gerisch (MPI for Biochemistry, Martinsried, Germany). The cells were grown at 21 °C in AX2 medium, supplemented with the antibiotic gentamicin at a concentration of 20 μg ml^{-1} (G-418, Biochrom AG, Berlin, Germany). Myosin II-null mutants were additionally supplemented with the antibiotic blasticidin at a concentration of 20 μg ml^{-1}. During cell culture in Petri dishes, the cell concentration was kept below 40%.

A benomyl ($C_{14}H_{18}N_4O_3$, Sigma-Aldrich, Germany) solution in phosphate buffered saline (PBS) was used as the microtubule depolymerization agent. The solution was added to the *D. discoideum* WT cell suspension 30 min before measurements at a concentration of 100 μM.

For microscopy experiments, cells were harvested from the Petri dishes by three successive washing steps with 17 mM K–Na phosphate buffered saline, adjusted to pH = 6.0 (PBS, Sigma Aldrich, Steinheim, Germany) and supplemented with 18 g l^{-1} maltose (Maltose monohydrate, Karl Roth GmbH, Karlsruhe, Germany). The cells suspended in the maltose–PBS solution were then added progressively to the Ibidi observation chamber (uncoated 8-well, Ibidi, Martinsried, Germany) and let to settle down, until a concentration of 5 to 10 cells per 125 by 125 μm (camera field of view) was achieved. This concentration ensures enough cells for statistics and few enough cells to exclude any cell–cell signaling, which would bias spontaneous migration.

Microscopy. The experimental chamber was mounted onto a Zeiss Axiovert 200M microscope and kept at $T = 21$ °C. Both brightfield and fluorescence images were acquired every 10 s with an exposure time of 150 ms by an Andor iXon Dv885 CCD camera (Andor, Belfast, UK), triggered by the corresponding AndorIQ software.

2. PDMS microstructures

All substrates used for migration assays were fabricated from transparent polydimethylsiloxane (PDMS) after casting, cross-linking and unpeeling from a silicon wafer-based master obtained by photolithography (Fig. 1G–I). These PDMS arrays of 10–12 μm high pillars of 4 μm diameter are arranged in a regular lattice or in a network of varying pillar density. Outside the pillar fields, large areas of the same chemical composition (PDMS) are used as a reference flat substrate.

Pillar structures and flat surfaces exhibit the same chemical composition (PDMS) throughout this study. In addition to that, both the wall and the base of the pillars, as well as the flat area between two pillars, exhibit homogeneous surface properties. Cells are not subject to any chemical attraction. Due to their low adhesivity, *D. discoideum* cell motility should not be strongly

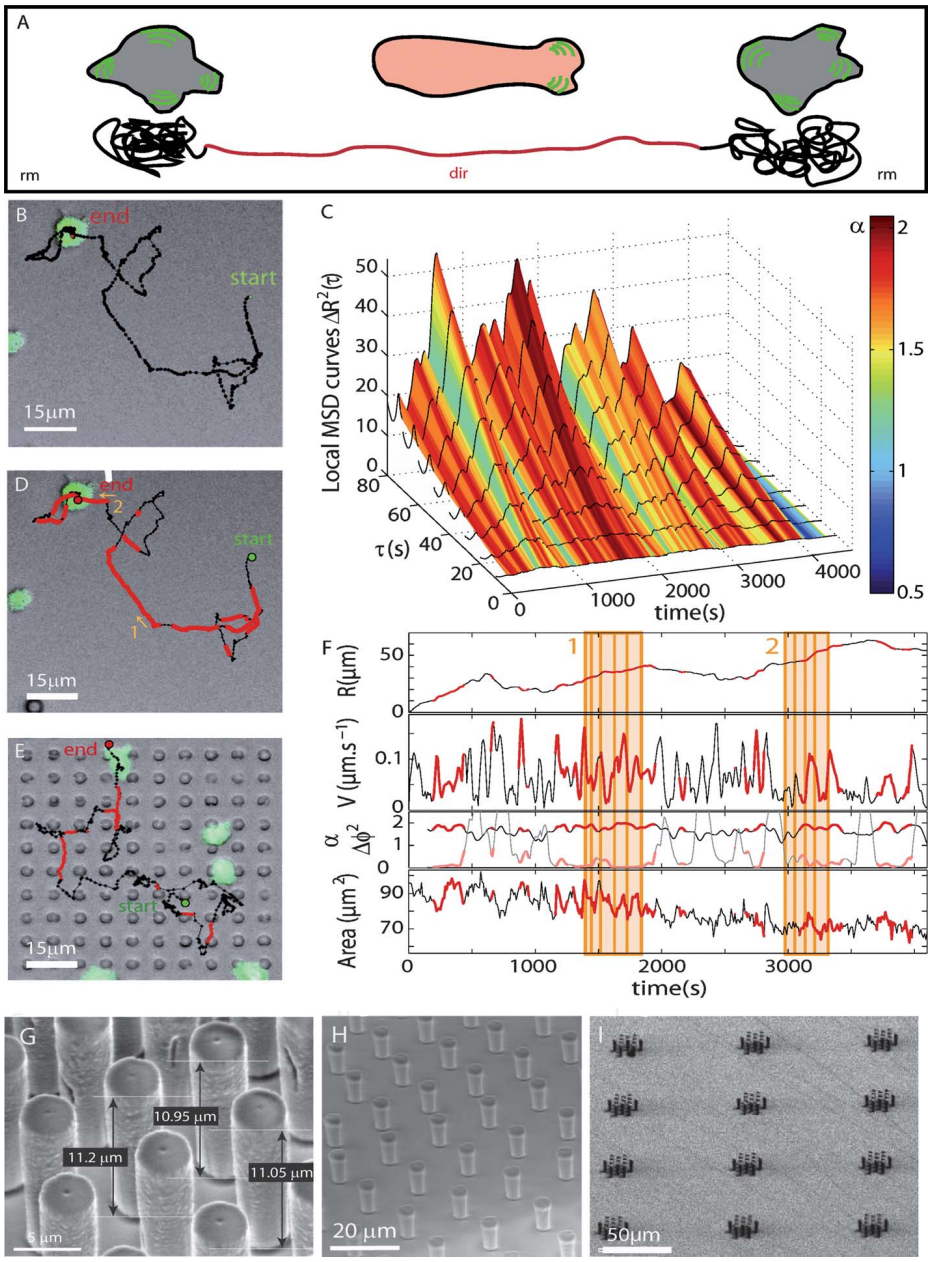

Fig. 1 Trajectory splitting into the two motion modes by the TRAnSpORT algorithm and PDMS microstructures. (A) Scheme of the transition from the random migration mode, the *rm*-mode (grey), where the cell exhibits multiple pseudopods, to a *dir*-run (red) characterized by one leading pseudopod (actin protrusions, represented in green). (B) Cell center-of-mass track on a flat surface before analysis. (C) Landscape of local MSD functions, calculated for each time t_i of the experiment (color-coded for α, the value of the power-law fitting exponent). The high α values, together with a high angular persistence, set the time points for which the motion is directed (*dir*-runs). (D and E) Cell track on flat PDMS (D) and within a rectangular micropillar array (E) after splitting into the 2 motion modes: *dir*-runs (red) and *rm*-modes (black). (F) Time course of the different variables defining the cell motion: cell position $R(t)$, instantaneous speed $V(t)$, local MSD exponent $\alpha(t)$ (black) together with the standard deviation of the angle variations $\Delta\phi^2(t)$ (grey) and area. The red portions of the curves indicate *dir*-runs. The windows highlighted in orange correspond to *dir*-runs (numbered 1 and 2 on image D). The speed oscillations are closely related to the area oscillations, emphasized by the orange lines. (G) SEM view (close up) of 4 µm diameter micropillars, with measured heights given after correction for the e-beam angle of 45 deg. (H–I) SEM micrographs of 4 µm diameter pillars, arranged in a regular square network (H), and in low-density arrays shaped as pillar islands (I).

influenced by surface chemistry, but the comparison between purely 2D and quasi 3D substrates on the same sample - consisting of the same material - ensures that any difference observed in cell motility modes can only be attributed to the well-defined micron-scale topography and not to different chemical properties of the surface.

Master fabrication. To obtain a master for PDMS molding, we used standard clean room microlithography procedures, and relied on the protocols presented by Steinberg *et al.*[26] A 3-inches silicon wafer (Si-Mat, Landsberg/Lech, Germany) was cleaned under nitrogen flow, covered with approximately 5 ml of a SU8-10 negative photoresist (Microchem, distributed by MicroResist

Technology, Berlin, Germany), and spin-coated at 3000 rpm for 30 s after progressive acceleration. A progressive soft bake (1 min at 65 °C and 2 min at 85–90 °C, with slow cooling down to room temperature) was followed by UV-illumination in a mask aligner (Süss MicroTec, Garching, Germany) through a Chromium mask (ML&C, Jena, Germany). This mask bears the structures transferred to the resist. Illumination times varied between 3 and 5 s according to the size of the desired structures. After a progressive hard bake (1 min at 65 °C and 2 min at 85–90 °C, with slow cooling down to room temperature), the unlit, and hence non-crosslinked photoresist, was washed away by two successive development baths in a SU8-specific developer (MicroResist Technology, Berlin, Germany). The photoresist structures obtained were stored at room temperature and exhibit stable properties for several months.

Preparation of the PDMS structures. To ensure proper PDMS unmolding, the resist master was silanized: vapour deposition of a fluorosilane (1*H*,1*H*,2*H*,2*H*-perfluorooctyltrichlorosilane, ABCR, Germany) was achieved under vacuum in one hour.

The PDMS base was mixed vigorously with the cross-linking agent at a 1 to 10 ratio (Sylgard 185 Silicon Elastomer Kit, Dow Corning, MI, USA). After degassing for 30 min under vacuum, 2–3 ml of PDMS were poured on the Si-photoresist master, and degassed again for 15 to 30 minutes under vacuum. PDMS cross-linking was obtained after 3 to 5 hours at 65 °C.

After cutting and peeling the PDMS structures off the master, the samples were exposed to Argon plasma for 30 s, to make the surface hydrophilic (Plasmaanlage "Femto", Diener Electronic, Nagold, Deutschland). They were then immediately transferred to the observation chambers (uncoated 8-wellplates, Ibidi, Germany), immersed in PBS solution, and kept sterile until further use with cells.

3. Cell tracking and trajectory analysis

All cell positions, areas and perimeters were retrieved from the acquired fluorescence images by a homemade plugin for the ImageJ analysis software (W. S. Rasband, U. S. National Institutes of Health, Bethesda, USA, http://imagej.nih.gov/ij/). The pillar positions were obtained from brightfield images with the same plugin. The cell trajectories were then processed by a homemade Matlab algorithm (The Mathworks Inc. Natick, USA), which is able to retrieve the local motion modes characterizing cell migration (Fig. 1C and F). Furthermore, the information retrieved on cell contact with pillars enables the study of the influence of micron-scale obstacles on *D. discoideum* motility modes.

Migration track retrieval by the *single cell tracker* plugin. Our single fluorescent cell tracking plugin has been adapted to follow the rapid motion of *D. discoideum* amoebae. This algorithm detects fluorescent cells as clusters of more than n bright pixels above an intensity threshold I_0 and in closer proximity than a distance ε. All three parameters (n, I_0, ε) can be freely adjusted to yield the most reliable cell clustering results. Tracking of each brightness cluster throughout the movie is achieved by finding, from one frame to the next, the most proximate cluster (in space and intensity), which is then identified as the same object. Each brightness cluster represents a single *D. discoideum* cell, and yields its center of mass position (X, Y), area (A) and perimeter (P) as a function of experiment time.

From brightfield images, the pillar positions can be determined employing the same plugin:¶ after image treatment (noise reduction) and thresholding, a precise map of the pillar outlines on each time frame is retrieved. This map is then used for determining and analyzing the cell-to-pillar contact (see "*Contact* evaluation"). Also, pillar positions were used to estimate a potential stage motion (long-term drift or short-term small oscillations) and to subtract it from the overall cell motion before analysis.

Cell motion analysis by the *TRAnSpORT* algorithm. Our *TRAnSpORT* (Time-Resolved Analysis for the Splitting Of Random Trajectories) routine[34–36] performs cell motion analysis, since it is capable of distinguishing two motility modes, one of them exhibiting high directional persistence and the other mode being an isotropic, diffusive-like random walk.

For each time point t_i, representing the center of a rolling window of $M = 30$ points, a local mean square displacement ΔR_i^2 (l-MSD, Fig. 1C) is calculated as a function of the time lag $\tau_k = k\delta t$ as follows:

$$\Delta R^2(t_i,\tau_k) = \Delta R_i^2(k) = \frac{1}{M-k+1}\sum_{j=-M/2}^{(M/2)-k}\left(\boldsymbol{R}\left(t_{i+j}+k\delta t\right)-\boldsymbol{R}\left(t_{i+j}\right)\right)^2 \quad (1)$$

where δt is the frame rate and $\boldsymbol{R}(t_i) = (X(t_i), Y(t_i))$ are coordinates of the center of mass. $T = M\delta t$ is the duration of the rolling window. We already showed that the resulting algorithm time resolution is of the order of $T/4$.[34] The l-MSD functions are then fitted by power laws $f_i(\tau) = A_i\tau^{\alpha_i}$, with the exponents α_i bearing information about the motility state at each point i (Fig. 1C and F).

The angle persistence function $\Delta\phi_i$ is calculated from the values of the velocity angle $\phi_i = (v_{x,i}, v_{y,i})$:

$$\Delta\phi(t_i,\tau_k) = \Delta\phi_i(k) = \sqrt{\frac{1}{M-k+1}\sum_{j=-M/2}^{(M/2)-k}\left(\phi\left(t_{i+j}+k\delta t\right)-\phi\left(t_{i+j}\right)\right)^2} \quad (2)$$

The l-MSD exponent and angle persistence values allow for a dissection of the trajectory into two different motility states: if α_i is close to 2 and $\Delta\phi_i$ is close to zero, the motion is classified as directed (or ballistic), and the cell is in a *dir*-run (see Fig. 1D–F, red phases). Otherwise, the motion is called random probing (or non-directed), and the cell is in an *rm*-mode. This is summarized by a binary probability for directed motion p_{dir}

$$p_{\mathrm{dir}} = \begin{cases} 1 & \text{if} \quad [2-\sigma_\alpha \le \alpha \le 2] \wedge [0 \le \Delta\phi \le \sigma_\phi] \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

The directed parts of the trajectory correspond to phases during which the amoeba crawls at a quasi-constant speed in a quasi-preserved direction. An angular deviation of $\sigma_\phi = 3\sigma_\alpha$ is

¶ Pillars appear as very bright disks, due to a light guiding effect along their axis.

allowed, so that *dir*-runs do not depend on slow changes in the motion direction (the value of σ_α is set to 0.3, to ensure correct discrimination between the *dir*- and *rm*-modes). Using both criteria, the exponent and angle persistence values are necessary due to the uncertainty of the α-value arising from the limited number of points in the rolling window.

Contact evaluation. The aim is to reliably measure the number of pillars that the cell is in contact with for each image. This is called the *total contact index* I_{cont} of a single cell during its trajectory. We extract dwell times, which are the periods for which the cell has been in contact with one or more pillars.

The probability for a cell c to be in contact with a pillar p in its vicinity depends on the distance between the cell and pillar, and more specifically, on the fraction of the cell area which is close to this pillar. More formally, we introduce a pairwise contact index $i_{\mathrm{cont,c-p}}$ which reflects this probability. In order to calculate this index, we define two parameters: ε (the critical cell distance or proximity in units of pixels) and N_{min} (the minimum cell area in units of pixels that has to be closer than ε to the pillar). If n_{cell}, the number of cell pixels in proximity to the pillar, is greater than N_{min}, then the index $i_{\mathrm{cont,c-p}}$ is set to 1. Otherwise, if n_{cell} is smaller than N_{min}, the index value is set to $n_{\mathrm{cell}}/N_{\mathrm{min}}$.

The total *contact* index for a given cell c in a given frame j is the sum over all pairwise indices corresponding to different pillars:

$$I^{c}_{\mathrm{cont}}(j) = \sum_{p} i_{\mathrm{cont},c-p}$$

This yields the total number of pillars the cell is in contact with. Running this analysis frame by frame allows for correlating the cell-to-pillar contact with the cell motility.

Through the use of the contact index, the statistics on the motion characteristics (instantaneous speed, α exponent and *dir*-run lifetimes) can be split into specific categories, corresponding to the phases during which the cell is in contact with none, one or two (or more) pillars. Such a distinction is crucial to understand the influence of topographical cues on cell migration.

Results and discussion

Cell migration analysis by a 2-state motility model

In this work, we investigate the influence of 3D environments on cellular migration modes. Using standard photolithography procedures and polymer molding, we fabricated arrays of pillars of 4 μm diameter and 10–12 μm height, made of transparent polydimethylsiloxane (Fig. 1G–I). Pillar structures and flat surfaces exhibit exactly the same chemical composition, ensuring that any observed effect is of topographical origin only. In order to compare the migration of *D. discoideum* cells on flat surfaces and in microstructured pillar fields, the recorded migration trajectories of fluorescent cells were analyzed with high spatial and temporal resolution by a two-state motility model.

First, image processing is used to capture the cell motion of GFP-labeled *D. discoideum* cells. The cell center of mass is identified and tracked throughout the movie. We then analyze individual cell migration trajectories by our local-MSD-algorithm: a predefined time window slides across the migration track and allows for the calculation of a local mean square displacement (MSD)[34] (see also Materials and methods). The migration mode is assigned as directed (*dir*-run) when two conditions are fulfilled: (i) the local MSD function is close to a quadratic law with the exponent α of the power-law fit close to 2, and (ii) the angular persistence of the trajectory is high, with a standard deviation $\Delta\varphi$ of the velocity angle close to 0. If these conditions are not fulfilled, the cell is in a diffusive-like random probing (*rm*-) mode (Fig. 1A–F). Each track is dissected into random probing states and directed runs.

Cell migration on a flat surface

The global MSD functions calculated over the entire trajectories yield information about the cell migration type as a whole. Typical example functions are given in Fig. 2N–Q (corresponding tracks shown in the inset). The motion is superdiffusive at short time scales and diffusive at long time scales with a typical cross-over time from one regime to the other. But this global analysis can only give a rough representation of cell migration as a random walk which, at short time scales, involves periods of rather directed migration. While the double fitting of the global MSD functions can only yield one characteristic cross-over time τ_C, our high resolution analysis of local motion types enables separation of the contributions of different migration modes, and analysis of the distribution of their precise characteristics: local MSD exponent, velocity and lifetime.

Spontaneous migration on a flat substrate is characterized by alternating phases of directed and random motion modes (Fig. 1A). Two directed runs (shown in red in Fig. 1D and E) are separated by a random phase during which the cell slows down, probes its environment and repolarizes. A new *dir*-run in another direction starts after this isotropic reorientational phase. After trajectory splitting by our local MSD analysis (Fig. 1C and F), one can retrieve the phase durations, together with specific migration parameters such as velocities of the *dir*-runs and diffusion coefficients for the *rm*-modes (Fig. 1F). The advantage of our time-resolved analysis lies within the fact that the parameters are only evaluated during the corresponding motion phases.

During migration on a 2D surface without external stimulus (see Fig. 1B and D, and Fig. 2B, F, J and N), WT cells show 35% of *dir*-runs, which are randomly distributed without preferred direction in the X–Y plane (Fig. 1D). Their lifetime decays over $\langle\tau_{\mathrm{d}}\rangle = 140$ s and the runs cover a distance of 11–21 μm, which corresponds to 1–2 cell lengths. The speed distribution function exhibits a Gaussian shape, characterized by a well-defined mean speed of $\langle V_{\mathrm{dir}}\rangle = 0.076\ \mu\mathrm{m\ s^{-1}}$ (averaged over all measured cells (Fig. 2B)). Each *dir*-run is composed of 2 to 4 cycles of expansion/retraction of the cell body, which also results in speed oscillations (Fig. 1F). By contrast, the *rm*-mode speed is significantly smaller ($\langle V_{\mathrm{rm}}\rangle = 0.057\ \mu\mathrm{m\ s^{-1}}$) and its distribution is best fitted by a log-normal function, indicating that numerous processes of various origins and time scales are involved here. Furthermore, *rm*-modes are typically concentrated within regions of less than 7 μm in diameter.

Cell migration modes in regular micropillar arrays

We applied the same analysis to cell migration of three different types of *D. discoideum* within a dense rectangular micropillar

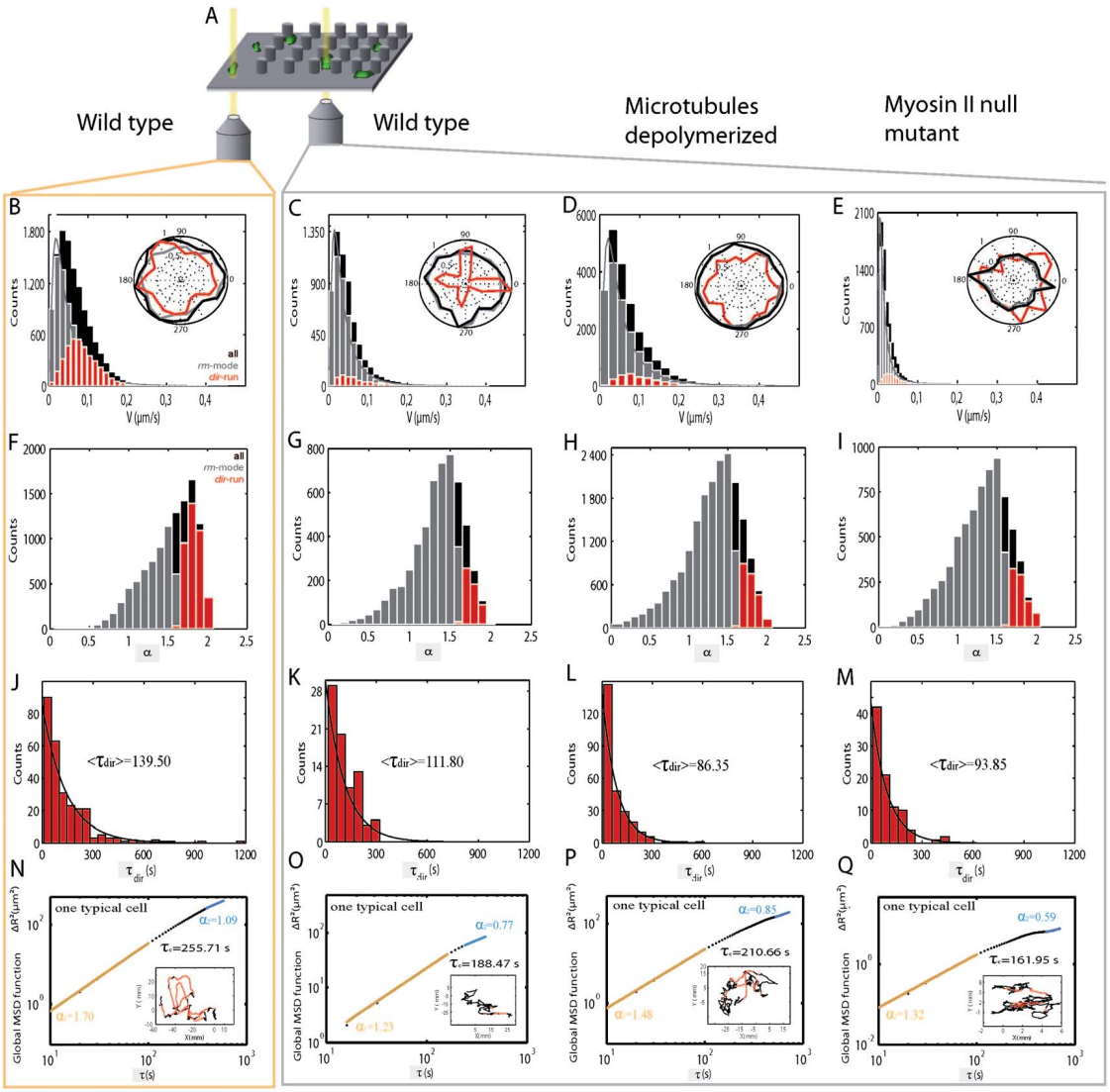

Fig. 2 Spontaneous amoeboid motility on a flat substrate *vs.* topography-guided migration within a square lattice of micropillars. (A) Scheme of the experimental setup: comparison of *D. discoideum* migration on flat and microstructured areas, both identical in chemical composition. (B–E) Instantaneous speed distributions, resulting from all analyzed cells on flat PDMS ($N = 27$ cells) and within this micropillar array wild type ($N = 13$ cells), benomyl treated cells ($N = 27$ cells) and myosin II-null mutants ($N = 14$ cells) (black: all data points, red: *dir*-runs, grey: *rm*-modes). Velocities during *rm*-modes are fitted by a log-normal distribution. Angle distributions of the velocity vector are shown in the insets (normalized to one). (F–I) Distributions of the exponent α, power-law fit of the local MSD functions, revealing the amount of *dir*-runs and *rm*-modes in the overall cell motion. (J–M) Distributions of the *dir*-run lifetimes, with fits by a simple exponential. (N–Q) Global MSD function, calculated on a single cell track shown in the inset: the short time/long time power law trends in the global MSD are highlighted in orange and blue respectively. The track has been split into *dir*-runs (red) and *rm*-modes (black).

network of 4 μm pillar-to-pillar distance, which is about one third of a typical cell diameter (Fig. 1E). Our measurements reveal a strong decrease in *dir*-run frequency compared to flat surfaces (Fig. 2, and Table 1): *dir*-runs represent only 10–12% of the cell motion (*vs.* 32% on flat). This is also reflected in a smaller average α value of the local MSD analysis.

In WT cells, we find a smaller average velocity $\langle V_{rm} \rangle$ of 0.043 μm s^{-1} in pillar structures as compared to 0.057 μm s^{-1} on flat substrates. However, the velocities during directed runs $\langle V_{dir} \rangle$ are larger (0.082 μm s^{-1} as compared to 0.076 μm s^{-1} on flat substrates). The most striking effect appears in the velocity angle distribution: while the *rm*-modes remain randomly oriented, the *dir*-runs reflect the network geometry (Fig. 2). So the presence of dense and regularly spaced micropillars heavily modifies *D. discoideum* migration modes: the decrease in frequency and duration of the *dir*-runs results in a much more confined cell motion. The rare directed runs are strongly guided along the lattice axes of the micropillar network, resulting in a biased run from pillar to pillar along these preferred directions (see ESI S2†).

Benomyl-treated cells, lacking microtubules, show in general larger velocities. Compared to WT cells in pillar fields, they do not stay in *dir*-runs as long as WT cells and exhibit rather random orientation. We find that the angle distribution of the *dir*-runs does not reflect the pillar geometry any longer.

Myosin-II null mutants show significantly reduced velocities in general. They are rather confined by neighbouring pillars due to

Table 1 Parameters describing the cell motility on a flat substrate *vs.* a square network of dense pillars (4 μm pillar-to-pillar distance), corresponding to Fig. 2

		Flat surface	Square network		
Micropillar substrate type		WT	WT	Benomyl treated	Myosin II-null
Number of cells in the statistics	N	27	13	27	14
Number of data points in the statistics	n	10 968	5315	19 467	7389
Exponent of the l-MSD power-law					
All	$\langle\alpha\rangle$	1.53	1.35	1.30	1.30
dir-runs	$\langle\alpha_d\rangle$	1.82	1.77	1.79	1.79
rm-modes	$\langle\alpha_r\rangle$	1.37	1.30	1.24	1.24
Instantaneous velocity/μm s^{-1}					
All	$\langle V\rangle$	0.059	0.043	0.061	0.019
dir-runs	$\langle V_d\rangle$	0.076	0.082	0.090	0.034
rm-modes	$\langle V_r\rangle$	0.057	0.043	0.060	0.018
dir-runs lifetime/s	$\langle\tau_d\rangle$	139.50	111.80	86.35	93.85

their extended cell size, as they cannot retract their tale effectively, and therefore do not show large displacements. Here, we find slight reflection of the pillar network geometry in the random migration angle distribution.

Migration modes at the flat/structured interface

In order to identify if significant trends in cell migration directed towards flat or structured surfaces can be revealed, we investigated motility modes and preferred cell localization at the interface between flat substrates and pillar fields. For that purpose, we designed low-density micropillar islands of about 50–100 μm in diameter, each consisting of 13 pillars. The islands are separated by wide flat areas (Fig. 1I). A *contact index*, *i.e.* the number of pillars a cell is in contact with along its trajectory (see Materials and methods), is defined. We find contact indices ranging from 0 to 4, depending on the inter-pillar distances. For cells migrating within a pillar island with several pillars in their vicinity, one subpopulation of cells moves from pillar to pillar with dwell times between 50 and 200 s, exhibiting mostly *dir*-runs (Fig. S3C† and 3E and F) whereas another cell population stays in contact with the same pillar for up to 500–900 s, mostly in the *rm*-mode (Fig. S3A†).

For inter-pillar distances larger than the typical cell diameter, where a single cell cannot touch two pillars at the same time, a striking "stick-and-go" type of motion is observed: the cells "jump" from pillar to pillar, exhibiting long *dir*-runs ($\langle l_d\rangle$ = 8.2 μm) thereby covering long distances (Fig. S1 and Movies S6 and S7†). This manifests the transition regime between 2D random migration and a migration that is modified by surface structures. In contrast, when the inter-pillar distance is slightly smaller than one cell diameter, cell motility is reduced by temporary trapping. Cells in contact with two pillars do not show any net displacement and cell locomotion can be stalled for durations of 900 s and longer (Fig. S2 and Movie S8†).

To quantify statistical dwell times, we averaged over the entire cell population (Fig. 3A). The dwell time distribution can be fitted by a double exponential $1-\beta\exp(-\tau/\tau_1)-(1-\beta)\exp(-\tau/\tau_2)$, which reflects the existence of two characteristic dwell times (fit calculated for the cumulative distribution). The main dwell time τ_1, representing around 77% of the touching events, is found to be ~74 s, and the second characteristic dwell time is $\tau_2 \approx 290$ s (corresponding to 23% of the touching events). This confirms that the vast majority of cells do not stay in contact for much more than one minute: most of the touching events are transient, along the "stick-and-go" motion of the cell. This proves that, in the case of low-density obstacles, cells are not stalled when they contact surface structures. Although pillars can be considered to be cell attractors, they should not be assigned as static traps. The dynamic trapping effect probably arises from feedback loops induced by internal signalling. When the cell is close to a pillar, pseudopods which are randomly protruded in all three dimensions during the *rm*-phase are more likely to touch the surface of 3D pillars than flat areas. Since it has been proven that a protrusion which touches a surface becomes a leading pseudopod with higher probability,[37] the likelihood for a cell to exit the random probing mode and to subsequently enter a new directed run, induced by a stabilized pseudopod, is increased in the vicinity of a micropillar (Fig. S3C and E†). However, the dwell time distribution includes values of up to 900 s, which reflects the rare but still observable very long touching events that were mentioned previously.

Cell ensemble partitioning between structured and unstructured surfaces

So far, we discussed migration trajectories of individual cells in the presence of pillars. We now take a different view and consider the steady-state situation of an ensemble of cells partitioning between pillar fields and unstructured areas. We assume that after a sufficient amount of time (typically 1–3 hours), a detailed balance of amoebae migrating in and out of the micropillar fields is reached. As shown in Fig. 3C, we perform a time and ensemble average for 27 cells in pillar structures for 1–3 hours and find that 22% of the cells are in contact with pillars. This percentage of cells in contact is now compared to the percentage of substrate in the vicinity of pillars (*i.e.* the area where cells can be in contact with a pillar) with respect to the total surface area. As indicated in Fig. 3B and C, we define the contact area by a circular region around a pillar of r = 7 μm in width, *i.e.* half the typical cell diameter of a *D. discoideum* cell. This area accounts for 8% of the total 2D surface: the "relocalization" factor is thus 0.22/0.08 = 2.75 (*cf.* ESI†). Accordingly, almost three times as many cells are in contact with pillars than would be expected for equal

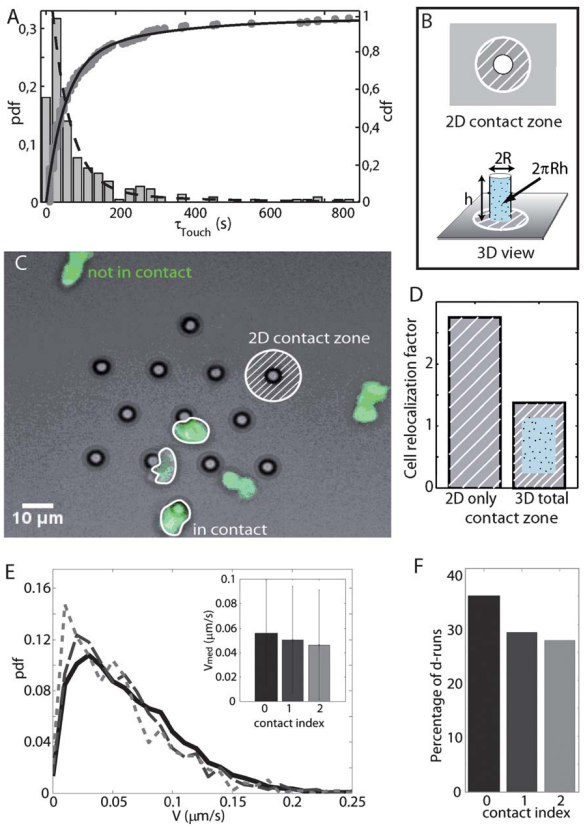

Fig. 3 Cell partition between the flat substrate and islands of micropillars. (A) Dwell times: probability distribution function (pdf, grey bars) and cumulative distribution function (cdf, grey dots) of dwell-times spent by cells in contact with pillars, for the whole population of cells migrating within and in-between islands of micropillars. The cdf is well fitted by a double exponential $1 - \beta\exp(-\tau/\tau_1) - (1 - \beta)\exp(-\tau/\tau_2)$ (black line) with characteristic times $\tau_1 = 74$ s and $\tau_2 = 289$ s, where $\beta = 0.77$ indicates that 77% of the touching events happen for a typical time τ_1. (B) Scheme of the 2D *vs.* 3D surface available around each pillar: (top) 2D projection of a pillar (radius $R = 2$ µm), with the hatched circle representing the zone within contacting distance for cells (radius $r = 7$ µm); (bottom) 3D view of the pillar with its walls offering an additional surface of $2\pi Rh$ (dotted blue). (C) Typical image underlining the cell relocalization effect: the zones within pillar-contacting distance (flat areas circumscribed around the pillars within which a cell can touch at least one pillar) represent only 25% of the 2D surface and yet, 3 out of 6 cells (50%) are in contact with pillars (white cell borders), the remaining 3 cells not being in contact with pillars (no cell borders indicated). This partition can only be explained by the 3D surface offered by the pillar walls, which yields an equivalent surface ratio of 50% in this specific case. (D) Cell relocalization factor, *i.e.* the percentage of cells in contact with pillars divided by the percentage of surface area within contact distance to pillars, averaged over all 27 measured cells. Without counting the surface available in the third dimension, this factor is almost 3. Taking the 2D contact zone *and* 3D surface offered by pillar walls into account, the total available surface in the vicinity of pillars increases drastically, which brings the relocalization factor close to 1: the cell distribution over the substrate exactly reflects the total available surface. (E) Velocity distribution and median velocity: compared to the reference on the flat surface (black, plain), the velocity distribution shows a slight shift towards low velocity values when the cells touch one (dashed, grey) or two pillars (dotted, light grey), which can also be seen in the median velocity values, shown in the inset (black: flat surface, grey: contact with 1 pillar, light grey: with 2 pillars). (F) Percentage of *dir*-runs: the frequency of *dir*-runs also decreases slightly with contact index. But on average, the cells keep moving directionally (more than 25% of the migration modes) and fast (more than 75% of the migration velocity on 2D), even when contacting 3D microstructures.

distribution according to the available 2D surface. However, if we additionally consider the vertical surfaces of the pillar walls (surface area: $2\pi Rh$), this situation changes drastically: in this case, the ratio of contact area ($2\pi Rh + \pi(R + r)^2$) to total area is 16%, which, considering experimental errors, corresponds to the percentage of cells in contact with pillars: the "relocalization" factor becomes $0.22/0.16 \approx 1$ (Fig. 3D). Hence we find that, on average, all cells are distributed equally with respect to the effective surface area within the three-dimensional microstructures. Given the length scales involved, it is natural to notice that cells do not distinguish between horizontal and vertical surfaces during adhesion and migration. We show that 2D projections, which are our familiar way of seeing things, can be misleading and even hide fundamental 3D-induced effects.

Conclusion

In summary, we investigated amoeboid migration on microstructured surfaces as a prototype example of migration in three-dimensional, natural environments. We found that the two-dimensional trajectories of *Dictyostelium discoideum* cells encountering micropillars show distinct differences from the basic two-state motility model on flat substrates.

Amoeboid migration modes are altered by the presence of 3D structures

According to the initial motility state, we discriminate two types of cell behavior upon pillar contact: randomly moving cells stay in contact with pillars, whereas fast moving cells in a directed run phase get deflected by the pillars. Cells migrating inside pillar arrays exhibit a characteristic stick-and-go behavior, if the pillar distance is large compared to the cell diameter. Furthermore, the partitioning of cells between pillar fields and flat areas allows us to conclude that cells distribute equally according to the total available surface in three dimensions.

Amoeboid cells migrate by maximizing contact with available surfaces

All our observations are consistent within the framework of a two-state migration model, where cells switch between phases of random pseudopod formation and phases where a single pseudopod is stabilized. Our experiments suggest that the switching from a randomly formed pseudopod into a stabilized pseudopod is enhanced by surface contact. This additional feature, in which the protrusion is reinforced independently of the spatial orientation of the contacted surface, guides cells through 3D environments.

In cells lacking microtubules or myosin II, we find significantly different behavior. Cells lacking microtubules do not show a pronounced attraction to pillars, although we find larger velocities during *dir*-runs. We conclude that microtubules enhance cellular ability to react with external 3D structures. In cells lacking myosin II, we find strongly decreased migration

velocities in both migration modes and a slight guiding behavior in the random mode due to the neighboring pillars.

All in all, the best frame to analyze our results on wild-type cells seems to be the model of contact-reinforced motility, or, more exactly, pseudopod stabilization upon surface contact with a subsequently increased probability to become the leading pseudopod. This leading pseudopod then sets the direction for a new *dir*-run, which, in the case of pillar contact, will be directed towards the pillar.

The effect of pillars, however, should always be considered as dynamic, as the overall cell motion remains a quasi-random walk. The emission of random protrusions is probably not biased by the presence of microstructures, while the *dir*-runs are. In the resulting mechanism of contact guidance, it is a stable cell polarization (the presence of a stable actin front at the leading edge due to surface-generated feedback loops and intracellular signalling) that leads to motional persistence towards microstructures. Amoeboid migration is hence guided by surface contact and is "haptotactic" in this sense. Considering the microstructure of soil, which is the natural habitat of *Dictyostelium discoideum* cells, contact-controlled motility is an advantageous strategy for amoebae to reach surface structures during their search for food, since bacteria preferably grow in dips and niches.

Possible advances for quantitative medical assays

As demonstrated here, experiments using micro-structured surfaces in combination with quantitative analysis of cell motility provide a useful tool to uncover details of cell–surface interactions and their coupling to cell motility.

In a medical context, this concept offers several opportunities for invasion assays, allowing for cell sorting on a large scale. By adding chemical gradients, which can be done in a very sophisticated way by now,[38] cells could be separated even more efficiently. The assay could easily be scaled up for insertion of whole cell sheets or even tissue parts to investigate collective cell motion phenomena. Related to this, the influence of edges and corners on cytoskeleton activity remains to be investigated in future experimental and theoretical work and should reveal crucial clues for the understanding of how topographical details of a substrate affect cell migration, as described for wall-like structures in ref. 39.

Acknowledgements

We thank E. Sackmann (TU München, Germany) and B. Lindner (Humboldt University, Berlin) for fruitful discussions, J. Rädler (LMU München, Germany) for helpful comments on the manuscript, S. Schulz (University of Heidelberg, Germany) for his kind help on the microlithography protocols, P. Altpeter and P. Paulitschke (LMU, Germany) for the SEM pictures. D.H. acknowledges funding from the Deutsche Forschungsgemeinschaft (grant HE5958-2-1) and from the Volkswagen-Foundation (grant I85100). C.L. acknowledges financial support by the Elite Network of Bavaria (International Doctorate Program NanoBioTechnology).

References

1 P. Friedl, *Curr. Opin. Cell Biol.*, 2004, **16**, 14–23.
2 J. Dalous, E. Burghardt, A. Muller-Taubenberger, F. Bruckert and G. Gerisch, *et al.*, *Biophys. J.*, 2008, **94**, 1063–1074.
3 C. M. Lo, H. B. Wang, M. Dembo and Y. L. Wang, *Biophys. J.*, 2000, **79**, 144–152.
4 P. N. Devreotes and S. H. Zigmond, *Annu. Rev. Cell Biol.*, 1988, **4**, 649–686.
5 C. A. Parent and P. N. Devreotes, *Science*, 1999, **284**, 765–770.
6 N. L. Jeon, H. Baskaran, S. K. W. Dertinger, G. M. Whitesides and L. Van De Water, *et al.*, *Nat. Biotechnol.*, 2002, **20**, 826–830.
7 C. L. Manahan, P. A. Iglesias, Y. Long and P. N. Devreotes, *Annu. Rev. Cell Dev. Biol.*, 2004, **20**, 223–253.
8 L. Song, S. M. Nadkarni, H. U. Bodeker, C. Beta and A. Bae, *et al.*, *Eur. J. Cell Biol.*, 2006, **85**, 981–989.
9 R. J. Petrie, A. D. Doyle and K. M. Yamada, *Nat. Rev. Mol. Cell Biol.*, 2009, **10**, 538–549.
10 D. R. Soll, D. Wessels, P. J. Heid and H. Zhang, *J. Muscle Res. Cell Motil.*, 2002, **23**, 659–672.
11 J. C. Del Alamo, R. Meili, B. Alonso-Latorre, J. Rodriguez-Rodriguez and A. Aliseda, *et al.*, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 13343–13348.
12 D. Heinrich, S. Youssef, B. Schroth-Diez, U. Engel and D. Aydin, *et al.*, *Cell Adhes. Migrat.*, 2008, **2**, 58–68.
13 Y. T. Maeda, J. Inose, M. Y. Matsuo, S. Iwaya and M. Sano, *PLoS One*, 2008, **3**, e3734.
14 Y. Iwadate and S. Yumura, *J. Cell Sci.*, 2008, **121**, 1314–1324.
15 L. Li, S. F. Norrelykke and E. C. Cox, *PLoS One*, 2008, **3**, e2093.
16 I. Weber, *Eur. J. Cell Biol.*, 2006, **85**, 915–924.
17 T. D. Pollard and G. G. Borisy, *Cell*, 2003, **112**, 453–465.
18 S. Benhamou, *Ecology*, 2007, **88**, 1962–1969.
19 H. C. Berg and D. A. Brown, *Nature*, 1972, **239**, 500–504.
20 P. Friedl, S. Borgmann and E. B. Brocker, *J. Leukocyte Biol.*, 2001, **70**, 491–509.
21 S. Cornillon, L. Gebbie, M. Benghezal, P. Nair and S. Keller, *et al.*, *EMBO Rep.*, 2006, **7**, 617–621.
22 E. Ponte, E. Bracco, J. Faix and S. Bozzaro, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 9360–9365.
23 J. Tan and W. M. Saltzman, *Biomaterials*, 2002, **23**, 3215–3225.
24 M. T. Frey, I. Y. Tsai, T. P. Russell, S. K. Hanks and Y. L. Wang, *Biophys. J.*, 2006, **90**, 3774–3782.
25 J. P. Kaiser, A. Reinmann and A. Bruinink, *Biomaterials*, 2006, **27**, 5230–5241.
26 T. Steinberg, S. Schulz, J. P. Spatz, N. Grabe and E. Mussig, *et al.*, *Nano Lett.*, 2007, **7**, 287–294.
27 M. Ghibaudo, L. Trichet, J. Le Digabel, A. Richert and P. Hersen, *et al.*, *Biophys. J.*, 2009, **97**, 357–368.
28 D. Wessels, H. Vawter-Hugart, J. Murray and D. R. Soll, *Cell Motil. Cytoskeleton*, 1994, **27**, 1–12.
29 G. Mahmud, C. J. Campbell, K. J. M. Bishop, Y. A. Komarova and O. Chaga, *et al.*, *Nat. Phys.*, 2009, **5**, 606–612.
30 S. Nagrath, L. V. Sequist, S. Maheswaran, D. W. Bell and D. Irimia, *et al.*, *Nature*, 2007, **450**, 1235–1239.
31 X. Trepat, M. R. Wasserman, T. E. Angelini, E. Millet and D. A. Weitz, *et al.*, *Nat. Phys.*, 2009, **5**, 426–430.
32 D. Gabriel, U. Hacker, J. Kohler, A. Muller-Taubenberger and J. M. Schwartz, *et al.*, *J. Cell Sci.*, 1999, **112**, 3995–4005.
33 D. Manstein, M. Titus, A. DeLozanne and J. Spudich, *EMBO J.*, 1989, **8**, 923–932.
34 D. Arcizet, B. Meier, E. Sackmann, J. O. Rädler and D. Heinrich, *Phys. Rev. Lett.*, 2008, **101**, 248103.
35 J. Mahowald, D. Arcizet and D. Heinrich, *ChemPhysChem*, 2009, **10**, 1559–1566.
36 C. Pelzl, D. Arcizet, G. Piontek, J. Schlegel and D. Heinrich, *ChemPhysChem*, 2009, **10**, 2884–2890.
37 D. Wessels, H. Vawter-Hugart, J. Murray and D. Soll, *Cell Motil. Cytoskeleton*, 1994, **27**, 1–12.
38 B. Meier, A. Zielinski, C. Weber, D. Arcizet, S. Youssef, T. Franosch, J. O. Rädler and D. Heinrich, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**(28), 11417–11422.
39 E. Sackmann, F. Keber and D. Heinrich, *Annu. Rev. Condens. Matter Phys.*, 2010, **1**, 257–276.

B Supplementary information for associated publications and manuscripts

B.1 Supporting information for associated publication P1

Single-cell mRNA transfection studies: Delivery, kinetics and statistics by numbers

By

Carolin Leonhardt, Gerlinde Schwake, Tobias R. Stögbauer, Susanne Rappl, Jan-Timm Kuhr, Thomas S. Ligon, Joachim O. Rädler

published in
Nanomedicine: *Nanotechnology, Biology, and Medicine*, vol.10, issue 4, feature article, 9 pages, 2014.
doi: 10.1016/j.nano.2013.11.008

Single-cell mRNA transfection studies: delivery, kinetics and statistics by numbers

Supplementary material

Carolin Leonhardt, Gerlinde Schwake, Tobias R. Stögbauer, Jan-Timm Kuhr, Susanne Rappl, Thomas S. Ligon and Joachim O. Rädler

Vector for *in vitro* Transcription

pEGFP is a commercial vector from BD Biosciences Clontech (4733bp, mRNA transcript 720 bases). It encodes a red-shifted variant of wild-type GFP, which has been optimized for brighter fluorescence and higher expression in mammalian cells. Sequences flanking eGFP have been converted to a Kozak consensus translation initiation site to further increase the translational efficiency in eukaryotic cells.

The pSTI-A120-vector (4746bp, mRNA transcript 1192 bases) is a friendly gift from Dr. Carsten Rudolph (LMU München, Kinderklinik und Kinderpoliklinik im Dr. von Haunerschen Kinderspital). We excised the luciferase gene from the vector pSTI-A120 by cutting with BamHI/EcoRI followed by a mutagenesis downstream from the T7 promoter to create a ribosome binding site.

The two vectors were chosen for the highest homology possible. However, they differ in their promoter because *in vitro* transcription of mRNA is only possible using the T7 promoter. To our best knowledge, the stabilizing β-UTR of the mRNA construct does not have an influence on transfection efficiencies. In addition to that, the sequences of the 3'UTRs of both vectors are comparable. The poly(A) tails of both transcripts, which have an impact on transcriptional stability, can be expected to be rather similar.

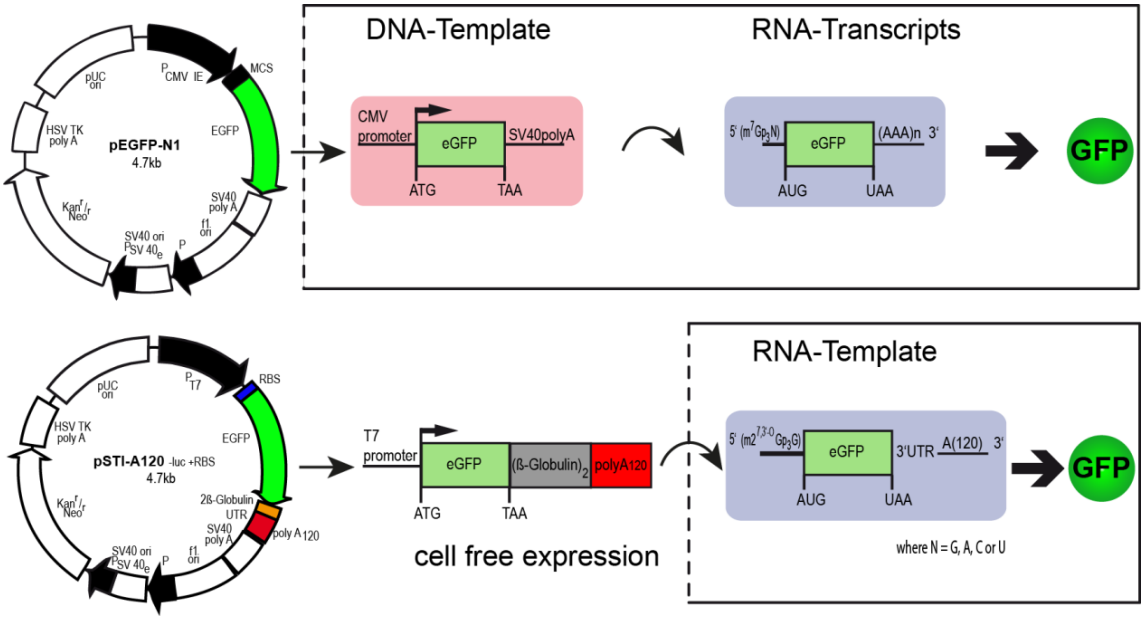

Figure S1. Vector maps of the pEGFP-vector and the pSTI-A120-vector. The analogous DNA- and RNA-templates are also shown.

Transfection for Single-Cell Time-Lapse Microscopy

A549 cells were grown in six-well plates for 24 hours before transfection from an initial seeding density of 1 x 10^5 cells/well for pDNA transfection and 5 x 10^4 cells/well for mRNA transfection, respectively. Equal weight amounts of pDNA and mRNA were used for transfection (20pg/cell). In the following, we will give the per-well-amounts of the reagents that we used for transfection per well: For the formation of lipoplexes, 2.5µl/4.75µl (mRNA/pDNA transfection) Lipofectamine were first diluted in 247.5µl/255.25µl OptiMEM transfection medium and 1.0µg mRNA/1.95µg pDNA were diluted in OptiMEM to add up to a total volume of 250µl. These mixtures were incubated at room temperature for 5 minutes. The Lipofectamine solution was then mixed with the nucleic acid solution, followed by another 20 minutes of incubation at room temperature. Immediately before transfection, the cells were washed with PBS and 500µl OptiMEM were added. Finally, 500µl of the lipoplex solution were added. The cells were incubated in a total transfection volume of 1ml at 37°C (5% CO_2 level) for one hour. The transfection medium was thereafter removed, the cells washed with PBS and re-incubated with Leibovitz's L-15 Medium containing 10% FBS right before monitoring of eGFP expression.

Transfection and Preparation for FC Measurements

A549 cells were grown in 24-well plates for 24 hours before transfection from an initial seeding density of 5 x 10^4 cells per well. Equimolar amounts of pDNA and mRNA were used for transfection (0.625pmol/well). In the following, we will give the amounts of the reagents that we used for transfection per well: For the formation of lipoplexes, 0.625µl/4.5µl (mRNA/pDNA transfection) of Lipofectamine was first diluted in 49.375µl/46.5µl OptiMEM transfection medium, respectively. Thereafter, 250ng mRNA/1.8µg pDNA were diluted in OptiMEM to add up to a total volume of 50µl. These mixtures were incubated at room temperature for five minutes. The Lipofectamine solution was then mixed with the nucleic acid solution, followed by another 20 minutes of incubation at room temperature. Immediately before transfection, the cells were washed with PBS and 150µl OptiMEM were added. Finally, 100µl of the lipoplex solution were added. The cells were incubated in a total transfection volume of 1ml at 37°C (5% CO_2 level) for one hour. The transfection medium was thereafter removed, the cells washed with PBS and re-incubated with full growth medium containing 10% FBS for 25 hours. The cells were detached from the 24-well plates with 150µl trypsin/EDTA per well. Subsequently, they were disaggregated by vigorous mixing with cold PBS containing 10% FBS and centrifuged for 5 minutes at 800 rpm. The cells were resuspended in 1ml ice-cold PBS and kept on ice until FC analysis.

Mathematical Modeling of eGFP Synthesis after mRNA Transfection

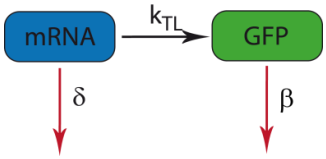

Figure S2. Kinetic rate model for GFP synthesis in mRNA-transfected cells.

We modeled eGFP expression of mRNA-transfected cells with a linear ansatz of two rate equations (Eq. **1** and **2** in the main text). Its solution (Eq. **3** in the main text) can be rewritten as follows:

$$G_{mRNA}(t) = \frac{k_{TL} m_0}{\beta - \delta}\left(e^{-\delta(t-t_0)} - e^{-\beta(t-t_0)}\right) \qquad \text{(S1)}$$

Equation (S1) shows that $G_{mRNA}(t)$ is symmetric with respect to the parameters β and δ, which represent the degradation rates of eGFP and mRNA, respectively. The values for β and δ are therefore interchangeable in our fitting formula. The eGFP degradation rate has been reported to be lower than that of mRNA [1]. This leads us to the assumption that the higher fitting value can always be assigned to δ whereas we believe the smaller value to represent β.
Another more illustrative representation of the (S1) is:

$$G_{mRNA}(t) = \frac{k_{TL} m_0}{\delta - \beta}\left(1 - e^{-(\delta-\beta)(t-t_0)}\right)e^{-\beta(t-t_0)} \qquad \text{(S2)}$$

This product can be divided into three terms, each describing one particular part of the eGFP time-courses: $G_{steadystate} = k_{TL} m_0 / \delta$ is the steady state protein level that a cell would reach if eGFP were not degraded ($\beta = 0$). The second term describes the monotonic increase of protein expression that we observe in all successfully transfected cells. The exponential decrease of the time-courses is defined by the GFP degradation rate β, which is the exponent of the last term. The slope of eq. (S2) at $t = 0$ corresponds to the initial eGFP synthesis rate in a transfected cell:

$$G'(0) = k_{TL} m_0 \qquad \text{(S3)}$$

For data analysis, we fitted each single time-course to the analytical solution (Eq. **S1/S2**). Here, the translation rate and the number of mRNA molecules were combined to $k_{TL}*m_0$ (i.e. the expression rate). All 281 single-cell time-courses were fitted in order to obtain distributions of the initial expression rate $k_{TL}*m_0$, the transfer time t_0, and the degradation rates β and δ. To validate this data analysis approach, we simulated mRNA transfection based on the Gillespie algorithm using the biochemical network simulator Copasi version 4.8[2]. The reactions that were taken into account were mRNA translation and mRNA as well as GFP degradation (see Figure S2). We used a translation rate of $k_{TL} = 170$/h, the degradation rates were assumed to be $\delta = 0.062$/h and $\beta = 0.056$/h. These are the mean values of the rate distributions that we obtained from fitting the experimental data. With these rates and an

initial number of 1000 mRNA molecules, we did 300 simulation runs (over a time span of 25 hours each) and fitted the resulting time-courses in the exact same way as we did for the experimental data. Figure S3 shows the resulting probability distributions of β and δ with the corresponding Gaussian fits:

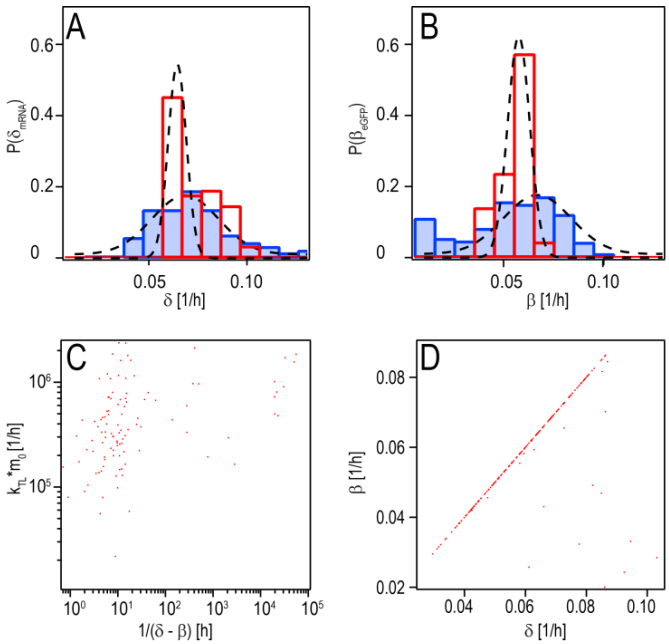

Figure S3. (*A*,*B*) Probability distributions of degradation rates (blue: experimental data, red: simulation, dashed lines: Gaussian fit curves). The portion of intrinsic noise (i.e. the width of the fit to the simulated data) contributing to the total noise (i.e. the width of the fit to the experimental data) is 26% for mRNA degradation and 34% for GFP degradation. (*C*, *D*) While the expression rate and the degradation rates are not correlated to each other (see Figure S3 *C*), a clear correlation between the degradation rates can be seen in Figure S3 *D*.

When we fit the distributions that we receive from simulation, the source of noise (i.e. the width of the Gaussian fit) can be of intrinsic nature only. This is due to the fact that our simulation is exclusively based on state transitions that are described by rates and does not include any cell-to-cell variations of these rates nor transfer noise. The variations of the rates that we extract from our simulated data are therefore a signature of the inherent stochasticity of the underlying biochemical processes. The additional width of the experimentally obtained rate distributions can be attributed to sources of extrinsic noise that are also present during transfection [3]. The intrinsic noise contributing a portion of about 30% of the overall noise seems to be a reasonable result and is in agreement with values that have been reported in literature [4].

Universality of mRNA Kinetics

The mRNA kinetics proves to be generic as three different cell types show the same mRNA expression curves. The single-cell expression time-courses are normalized with regard to the expression maximum and the onset time t_0. For comparison, the inserts show the original time-courses. The normalized expression curves nearly collapse onto a single expression master curve, which is fitted by Eq. **3** of the main text. The variance in the single-cell courses is due to cell-to-cell variability in the mRNA and eGFP degradation rates.

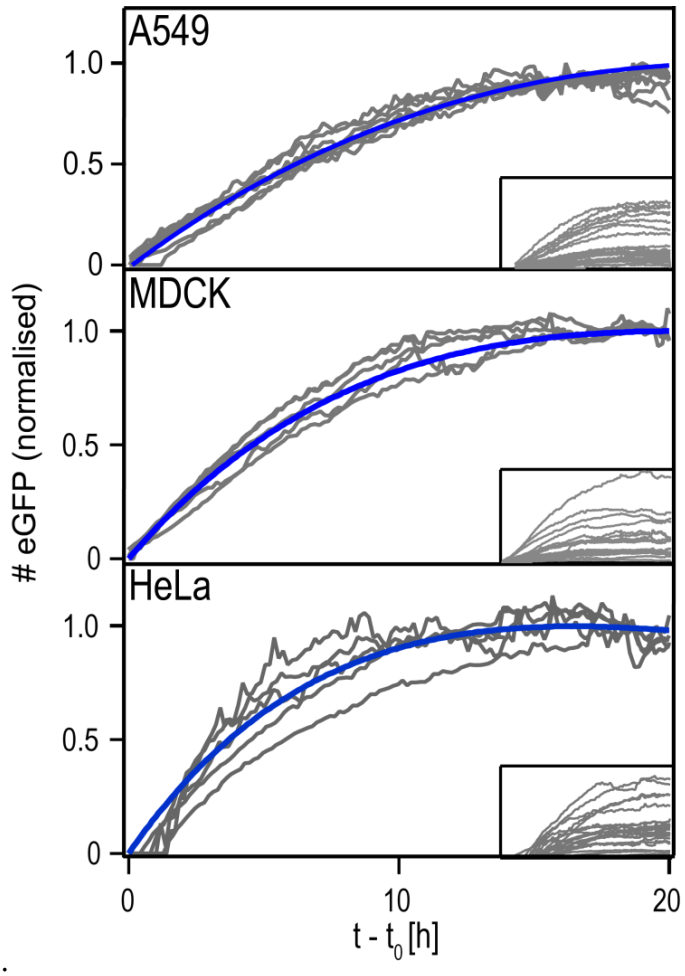

Figure S4. Normalized mRNA expression time-courses for various cell lines. Exemplary time-courses were normalized to their maximal value and shifted by the fitted onset time. Blue lines show the resulting master curve, raw data are shown in inserts. Largely cell-type independent behavior is observed.

A Stochastic mRNA Delivery Model

In order to describe the process of transfection by numbers, we developed a stochastic mRNA delivery model. During incubation, a portion of the lipoplexes in the transfection medium settles down onto the cell membrane. Some cells incorporate a fraction of the lipoplexes that are attached to their membrane by the formation of endosomes around lipoplexes. After uptake, some of the endosomes inside a cell lyse and part of their payload is successfully unpacked so that translation of the transfected mRNA can start. We will now describe in more detail how we developed a mathematical expression for the probability to find a certain amount of lipoplexes inside a cell:

Let η be the number of endosomes. The number of lipoplexes C inside a single endosome ($\eta = 1$) is assumed to be given by:

$$P_1(C) = \frac{L_{eff}^C}{C!} e^{-L_{eff}} \quad , \qquad \text{(S4)}$$

which is a Poisson distribution with mean L_{eff}. However, in a given cell the number of lysed endosomes is not necessarily $\eta = 1$ but η is rather a small, random number itself and therefore also Poisson distributed (with a mean N_{eff}). The probability of finding η endosomes in a cell then is

$$P_2(\eta) = \frac{N_{eff}^\eta}{\eta!} e^{-N_{eff}} \qquad \text{(S5)}$$

The probability to have C lipoplexes in a cell (not per endosome) can be written as

$$P_3(C) = \sum_\eta^\infty P_4(C|\eta) P_2(\eta) \qquad \text{(S6)},$$

summing up all possible ways of how C lipoplexes can be delivered to a cell (e.g. in only one single endosome $\eta = 1$, in two endosomes $\eta = 2$, etc.). We now need to find an expression for the probability P_4 $(C|\eta)$ to have C lipoplexes in a cell, given that η endosomes have lysed inside this particular cell. As the endosomes are statistically independent of each other, we can convolve $P_1(C)$ with itself η times. Here, we take advantage of the fact that the convolution of two Poissonians with means λ_1 and λ_2 is itself a Poissonian with mean $\lambda_1 + \lambda_2$.

The convolution of η Poissonians (each with mean L_{eff}) hence is a Poissonian with mean ηL_{eff}:

$$P_4(C|\eta) = \frac{(L_{eff}\eta)^C}{C!} e^{-L_{eff}\eta} \qquad \text{(S7)}.$$

The overall probability to find C lipoplexes in a given cell then is

$$P_3(C) = \sum_{\eta=0}^{\infty} P_4(C|\eta)P_2(\eta) = \sum_{\eta=0}^{\infty} \frac{(L_{eff}\eta)^C}{C!} e^{-L_{eff}\eta} \frac{N_{eff}^{\eta}}{\eta!} e^{-N_{eff}} \quad \text{(S8)}.$$

As every cell that contains one or more lipoplexes is considered to be successfully transfected, we can calculate the transfection efficiency by summing $P_3(C)$ from $C = 1$ to infinity:

$$P(C \geq 1) = \sum_{C=1}^{\infty} P_3(C) = \sum_{\eta=0}^{\infty} \frac{N_{eff}^{\eta}}{\eta!} e^{-N_{eff}} \sum_{C=1}^{\infty} \frac{(L_{eff}\eta)^C}{C!} e^{-L_{eff}\eta} \quad \text{(S9)}$$

Here, we changed the order of summation. The second sum is a sum over a Poissonian excepting the term for $C = 0$. Since all Poissonians are normalized, this sum gives

$$1 - e^{-L_{eff}\eta}.$$

Finally, we get an expression for the transfection efficiency:

$$P(C \geq 1) = \sum_{\eta=0}^{\infty} \frac{N_{eff}^{\eta}}{\eta!} e^{-N_{eff}} \left(1 - e^{-L_{eff}\eta}\right) = 1 - exp\left\{N_{eff}\left(e^{-L_{eff}} - 1\right)\right\} \text{(S10)}.$$

To fit the dose-response curve, we assumed L_{eff} to be proportional to the mRNA concentration c in µg ($L_{eff} = \lambda \times c$) and used this as fit parameter in order to extract N_{eff} from a fit of Eq. **S10** to the dose-response curve.

Estimation of Lipoplex Size and mRNA Load

To estimate the number of mRNA molecules per lipoplex, <*m*>, we measured the hydrodynamic radius of lipoplexes using fluorescence correlation spectroscopy and calculated <*m*> under the assumption of charge neutralization between the negatively charged nucleotides and the positively charged lipid analogues. Lipoplexes exhibit a configuration of intercalating mRNA and lipid-bilayer planes [5]. They can be approximated as cubes of an edge length that is twice the hydrodynamic radius of a lipoplex. A phospholipid's head group that is incorporated into a bilayer exhibits an average surface area of 70Å^2. The repeat distance from layer to layer is 65Å [5]. Lipofectamine is a 3:1 solution of DOSPA (four positive charges) and DOPE (neutral). To be more precise, we used Lipofectamine2000®, which we assume to be composed as Lipofectamine for this estimation. The N/P ratio in all single-cell experiments was about 5.4 and the mRNA we used consists of 1192 nucleotides. Taken together, this leads to an average number of

$$< m >= \frac{3 \cdot (120nm)^3 \cdot 2}{70 \cdot 1192 \cdot 5{,}4 \cdot 65\,\mathring{A}^3} \approx 350 \quad \text{(S11)}$$

mRNA molecules per lipoplex.

Estimation of the Fraction of Lipoplexes Adsorbed by the Cells

After one hour of incubation, the number of lipoplexes per cell is on the order of six in our experiments (see Figure 6*D* in the main text). In this section, we estimate the portion of all administered lipoplexes that is delivered toward a cellular monolayer within one hour incubation time. The transfer is diffusion limited because the theoretical sedimentation velocity of lipoplexes is negligible (about 0.4 nm/s).

The diffusion length is given by

$$\xi_{diff} = \sqrt{2Dt} = \sqrt{\frac{k_B T t}{3\pi\mu R}} \qquad \text{(S12)}$$

With k_B = 1.38x10^{-23}J/K, T = 310K, R_{Lipoplex} = 60nm, μ = 0.7x10^{-3}Pa*s and t = 1hour follows ζ_{diff} = 200µm. For an mRNA dose of 1µg and the derived package size of 350 mRNA molecules per lipoplex, we determined a concentration of 9x10^9 lipoplexes per ml transfection medium using FCS. At this lipoplex concentration, about 3000 lipoplexes should settle down on each cell within one hour if we assume an average cell area of 1600µm^2. However, this number is based on a "stick-and-hit" assumption where every lipoplex that hits the cell membrane stays there. In our experiment, we find no more than 15 lipoplexes per cell even after five hours of incubation at the above discussed dose (see also Figure 6*D*). Such small amounts of successfully delivered lipoplexes and hence the big loss of material must be caused by unknown side-effects and remains to be studied in more detail.

Determination of Gene Expression Onset Time after pDNA Transfection

In order to determine the onset time after pDNA transfection, the heuristic function $I(t) = \frac{I_{\max}}{2}\left[1 + \tanh\left(\frac{t - t_{1/2}}{t_0}\right)\right]$ was fitted to the recorded fluorescence time-courses. The time of expression onset, t_0, was calculated from the time of the half-maximal expression $t_{1/2}$ and the slope at that point as has been described in[6].

Calibration of Quantitative Fluorescence Imaging (QFI)

For quantification of eGFP expression levels in terms of eGFP molecules per cell, we imaged a polydimethylsiloxane (PDMS) micro-channel system filled with an eGFP standard solution at exact same microscopy and camera settings. The calibration system consisted of five channels. Each of them was 20µm wide, 8µm high, and 10mm long. The PDMS cast was made from a silicon master using standard photolithography. The PDMS cast was mounted on a glass cover slip via oxygen-plasma hydrophilisation. In order to prevent eGFP adsorption, the channels were passivated by incubation with PBS buffer containing 3µM Pluronic F108 for 20 hours at 4°C. The channels were consequently washed with PBS three times and dried under nitrogen flow.

The eGFP calibration standard is a His-tagged version of eGFP based on the same plasmid that was used for our single-cell transfection experiments. His-tagged eGFP was expressed in E. coli and purified using High Performance Liquid Bioaffinity Chromatography. The eGFP stock solution contained PBS (pH 7.4), Tween (0.001%, to prevent binding to surfaces), and sodium azide (0.002%, as a preservative). In the range of the stock solution's pH value, eGFP fluorescence is expected to be pH insensitive [7]. The eGFP concentration of the stock solution was determined photometrically using a NanoDrop 1000 spectrophotometer (Thermo Scientific, Germany) and independently checked by fluorescence correlation spectroscopy (FCS) on a ConfoCor2 (Zeiss, Germany).

The five channels of the calibration chip were filled with a concentration serial dilution of the calibration standard. A z-stack of the channels was recorded using the same 10x objective (focal depth of 8.5µm), exposure-time, and binning as used for the single-cell experiments. Calibration images were corrected for uneven illumination and auto fluorescence effects. In the calibration curve (Figure S6 *B*) the mean grey values (counts per pixel) were assigned to the calculated numbers of eGFP molecules in the corresponding channel volume knowing the imaged area per pixel and the height of the channels.

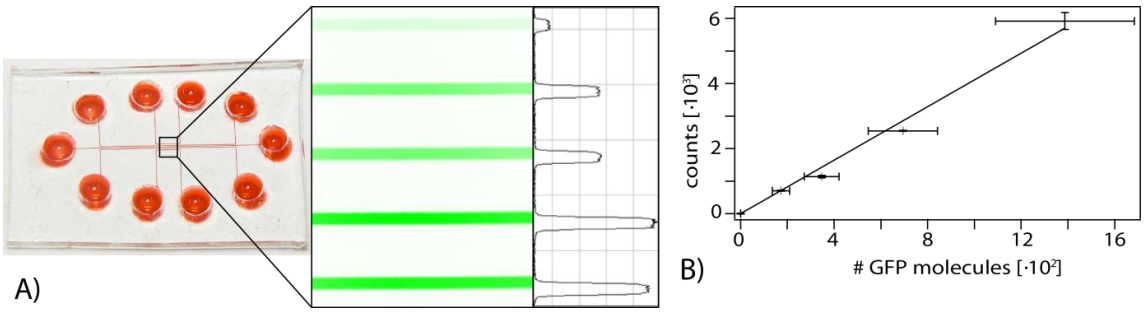

Figure S6. (*A*) PDMS micro channels were filled with a dilution series of eGFP solution and used to calibrate the grey values of microscopy images of eGFP expressing cells to numbers of eGFP molecules per cell. Left: PDMS chip with reservoirs and channels (filled with food colouring here for better visibility). Right: enlarged micrograph of the eGFP-filled channels. The second and third as well as the fourth and fifth channel (from the top) contain the same eGFP concentration, respectively. The far-right graph shows the corresponding brightness line plot. (*B*) Calibration curve for the conversion of grey values to numbers of eGFP molecules.

Rapid Reporter Turnover Using a Destabilized eGFP Variant

As we mentioned in the main text, the use of the destabilized eGFP variant, namely pd2EGFP-N1, as a fluorescent reporter circumvents the problems caused by longer observation periods. Figure S7 shows two exemplary curves of A549 cells that were transfected with mRNA encoding for destabilized eGFP. Here, the elimination phase of mRNA can be seen within 25 hours.

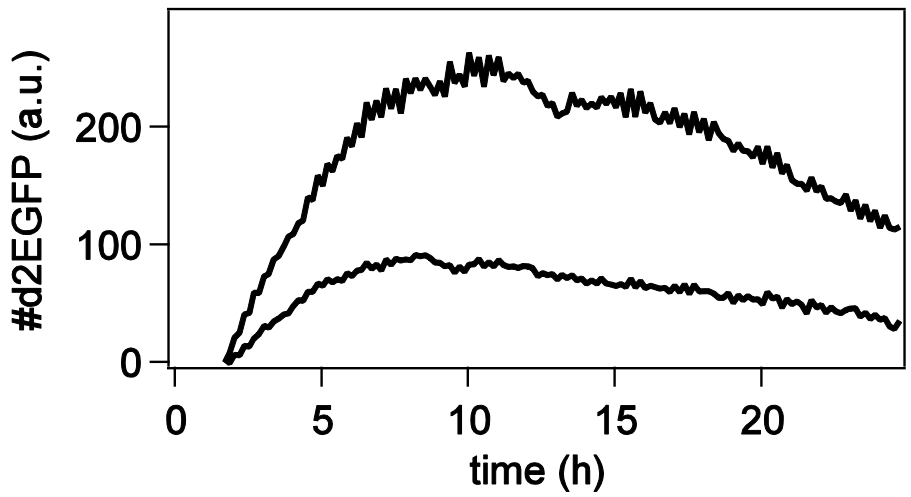

Figure S7. Two exemplary curves showing the time-course of d2eGFP expression after transfection of A549 cells with a destabilized eGFP variant. A decrease of the protein level to half the maximum expression can already be seen after about 20 hours in both cases.

References

1 Corish, P. & Tyler-Smith, C. Attenuation of green fluorescent protein half-life in mammalian cells. *Protein Engineering* **12**, 1035-1040, doi:10.1093/protein/12.12.1035 (1999).

2 Hoops, S. *et al.* COPASI—a COmplex PAthway SImulator. *Bioinformatics* **22**, 3067-3074, doi:10.1093/bioinformatics/btl485 (2006).

3 Blake, W. J., Kaern, M., Cantor, C. R. & Collins, J. J. Noise in eukaryotic gene expression. *Nature* **422**, 633-637, doi:http://www.nature.com/nature/journal/v422/n6932/suppinfo/nature01546_S1.html (2003).

4 Swain, P. S., Elowitz, M. B. & Siggia, E. D. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proceedings of the National Academy of Sciences* **99**, 12795-12800, doi:10.1073/pnas.162041399 (2002).

5 Rädler, J. O., Koltover, I., Salditt, T. & Safinya, C. R. Structure of DNA-Cationic Liposome Complexes: DNA Intercalation in Multilamellar Membranes in Distinct Interhelical Packing Regimes. *Science* **275**, 810-814, doi:10.1126/science.275.5301.810 (1997).

6 Schwake, G. *et al.* Predictive modeling of non-viral gene transfer. *Biotechnology and bioengineering* **105**, 805-813 (2009).

7 Karasawa, S., Araki, T., Yamamoto-Hino, M. & Miyawaki, A. A Green-emitting Fluorescent Protein from Galaxeidae Coral and Its Monomeric Version for Use in Fluorescent Labeling. *Journal of Biological Chemistry* **278**, 34167-34171, doi:10.1074/jbc.M304063200 (2003).

B.2 Supporting information for publication P3

Contact-controlled amoeboid motility induces dynamic cell trapping in 3D-microstructured surfaces

By

Delphine Arcizet, Sofia Capito, Mari Gorelashvili, Carolin Leonhardt, Marion Vollmer, Simon Youssef, Susanne Rappl, Doris Heinrich

published in
Soft Matter 2012,8,1473-1481.
doi:10.1039/c1sm05615h

Electronic Supplementary Material (ESI) for Soft Matter

Contact-controlled amoeboid motility induces dynamic cell trapping in 3D-microstructured surfaces

Delphine Arcizet[◊¤], Sofia Capito[◊], Mari Gorelashvili, Carolin Leonhardt, Marion Vollmer, Simon Youssef Susanne Rappl, and Doris Heinrich[*]

Center for Nanoscience (CeNS) and Faculty of Physics, Ludwig-Maximilians-Universität München, Geschwister-Scholl-Platz 1, 80539 München, Germany

[*] To whom correspondence should be addressed; E-mail: doris.heinrich@lmu.de
[◊] These two authors contributed equally to this work.
[¤] Current address: Institut de Biologie Structurale Jean-Pierre Ebel, Commissariat à l'Energie Atomique, 41 rue Jules Horowitz, 38027 Grenoble

MULTIMEDIA FILES AND SUPPORTING INFORMATION

This file contains:

A. Supplementary Figures and Tables

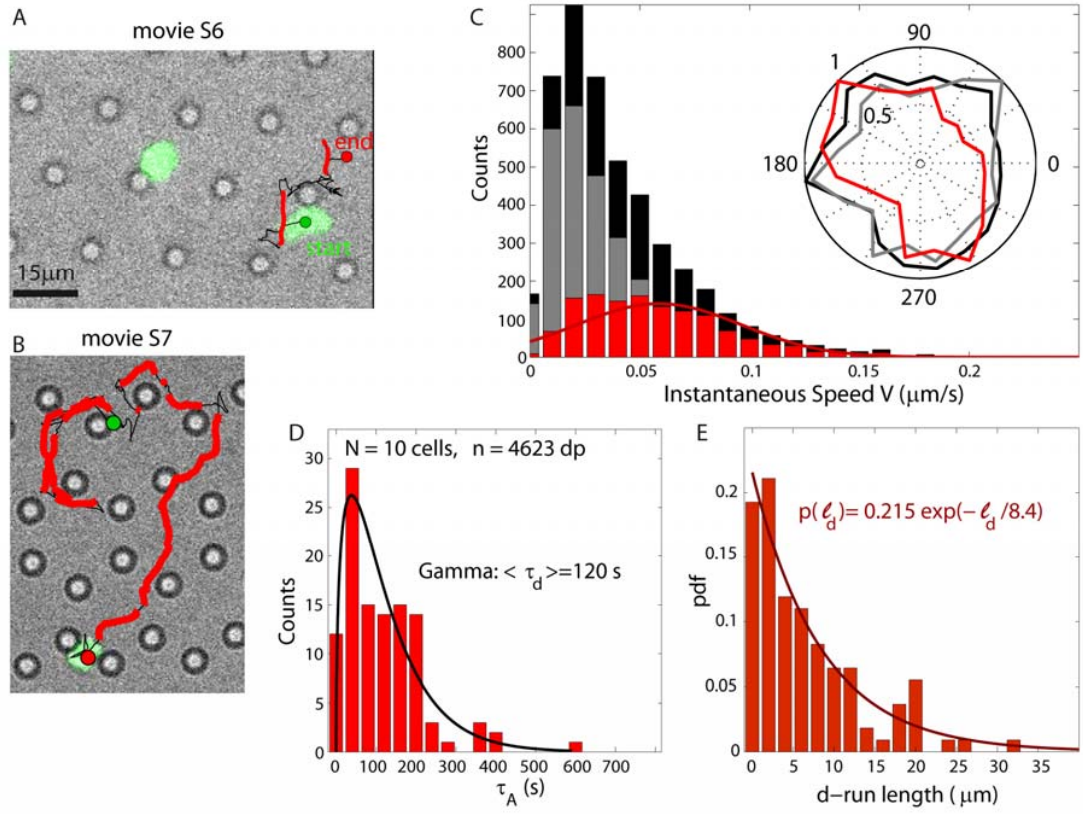

Figure S1: *D. discoideum* migration within pillar fields of high interpillar distance ($>15\,\mu m$).
(A – movie S6, B – movie S7) Typical cell migration tracks, displaying *dir*-runs from pillar to pillar: the cell (green) is efficiently guided from pillar to pillar, and covers a very long distance (B).
(C-E) Statistics on 10 measured cells. (C) The speed distribution is shifted towards lower values compared to flat, and the angle distribution remains random for *dir*- and *rm*-modes (insert). (D) The *dir*-run lifetimes are fitted by a Gamma distribution of average value 120 s, 22% smaller than on flat. (E) The distribution of distances covered during *dir*-runs is fitted by an exponential decay of typical cut-off length 8.4µm.

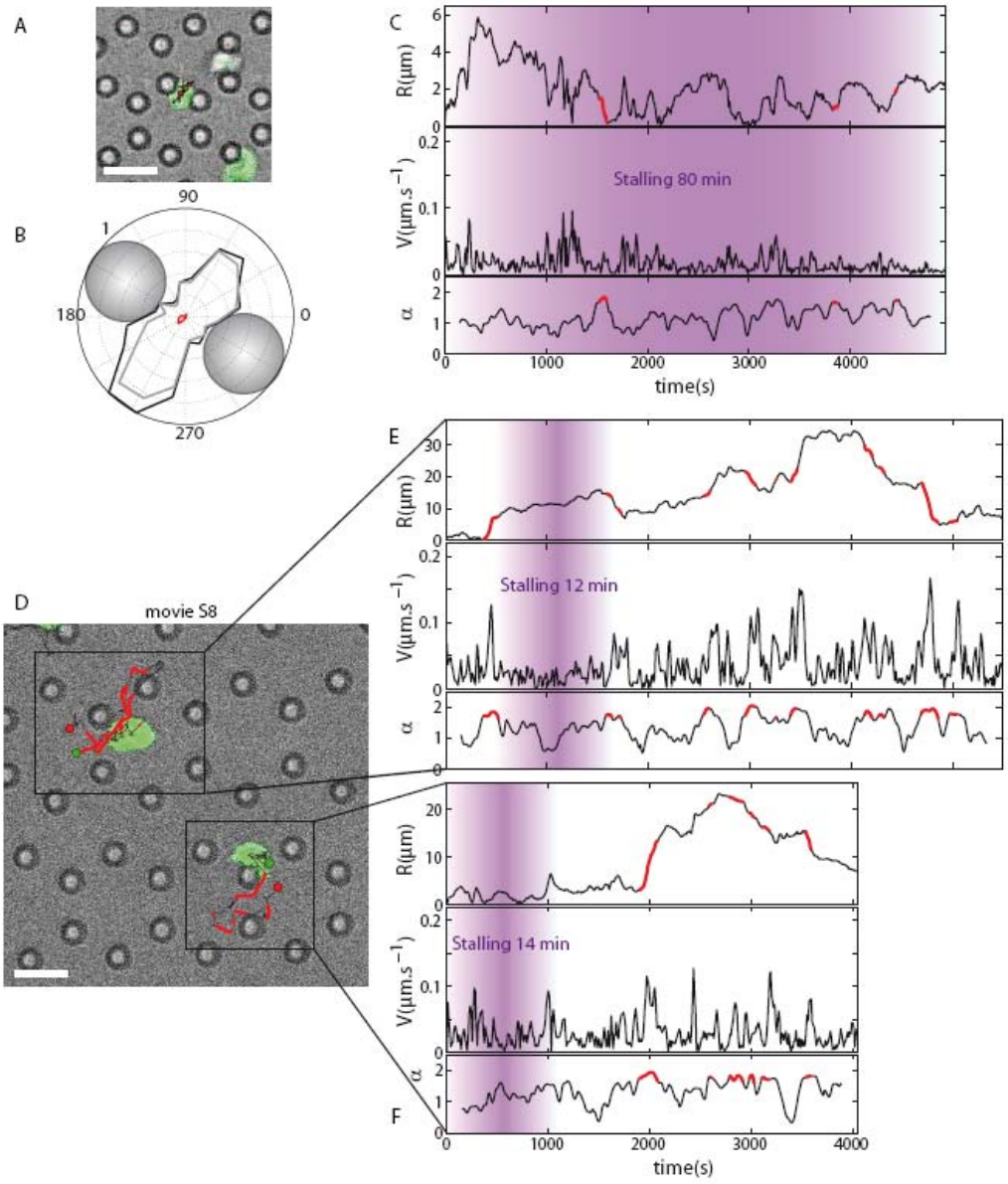

Figure S2: *D. discoideum* cells stalling within a pillar field where the interpillar distance equals the cell diameter.
(A-C) A cell is stalled between two pillars, during the whole recording time of over 80 min (A). The angle distribution shows back and forth oscillations between pillars without any net motion (B). The time series reveals a strong decrease in cell velocity (C). (D-F- movie S8) Two cells are stalled for a certain fraction of their recorded trajectory, as evidenced through their tracks (D), and time series (E, F).

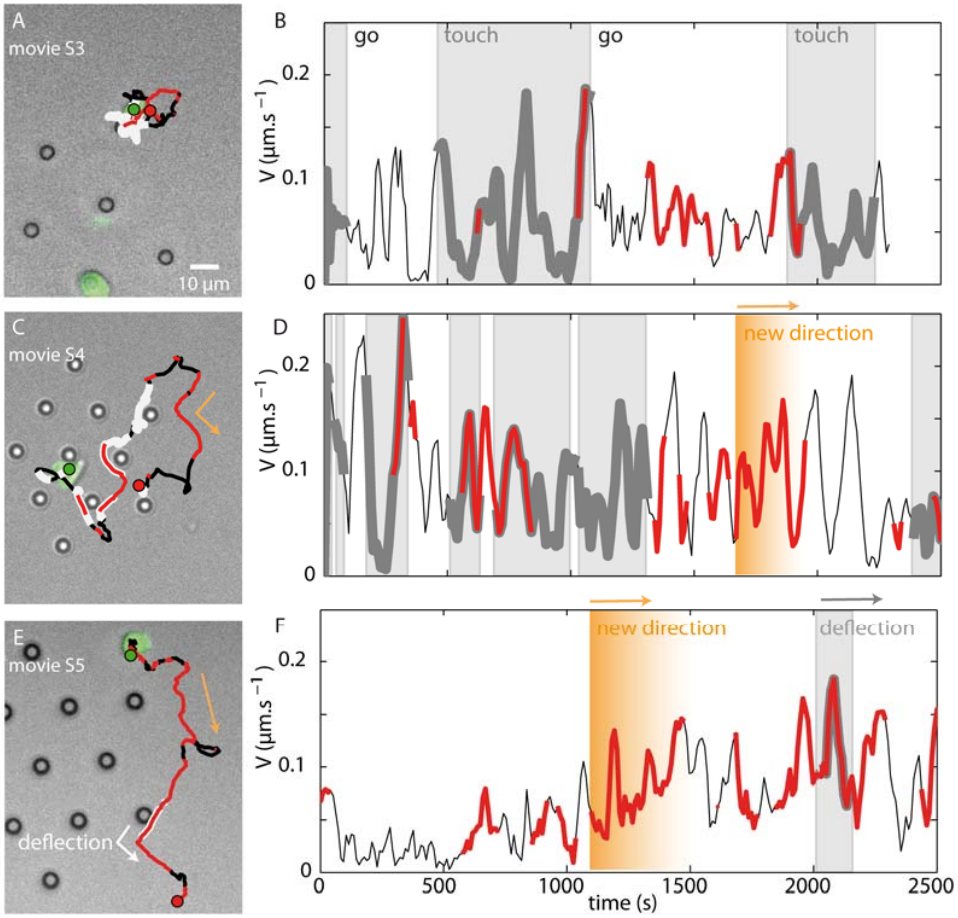

Figure S3: Modification of spontaneous migration upon contact with micron-scaled surface features.

Various scenarios of cell-pillar interactions: micrographs with individual cell tracks (A, C, E, with "touch" phases in white, *dir*-runs in red) and time sequences of the instantaneous cell speed *V* (B, D, F, "touch" phases in light gray).

(A, B) Revisiting track: the cell goes back and forth to the same pillar in an *rm*-type of motion for 40 min. Very rare *dir*-runs are present when the cell tries to escape from the pillar, and the cell is essentially trapped around the pillar.

(C, D) Guided track: the cell is guided from pillar to pillar within an island of microstructures, exhibiting a "stick-and-go" type of motion: after displaying several sequences of touching (grey) and going, the cell leaves the island, probes its surrounding flat environment, and finally repolarizes into a new direction (orange) until it finds a pillar again.

(E, F) Deflecting track: the cell probes its surrounding flat environment, polarizes into a certain direction (orange shading), increasing its speed (red arrow). After a new probing and re-polarizing event, V increases again (red arrow) and upon the next pillar contact, the cell is simply deflected by the pillar.

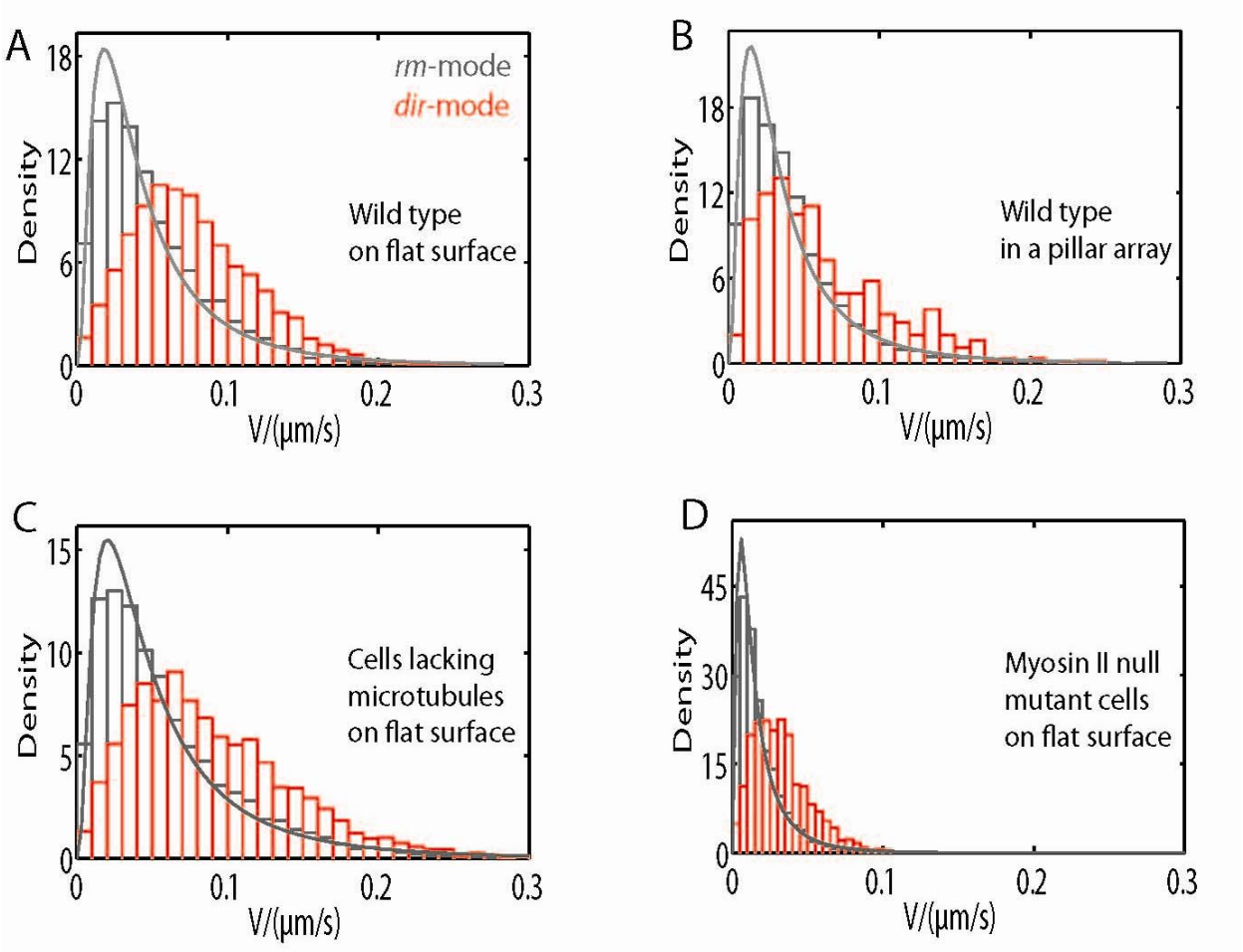

Figure S4: Quantitative Analysis of Velocity Distributions

(A) Instantaneous speed distributions, resulting from all analyzed wild-type cells on flat PDMS (red: *d*-runs, grey: *rm*-modes). The speed distribution of *rm*-modes is best fitted by a log-normal curve. (B) Instantaneous speed distributions, resulting from all analyzed wild-type cells within a micropillar array (red: *d*-runs, grey: *rm*-modes). The speed distribution of *rm*-modes is best fitted by a log-normal curve. (C) Instantaneous speed distributions, resulting from all analyzed cells lacking microtubules within a micropillar array (red: *d*-runs, grey: *rm*-modes). The speed distribution of *rm*-modes is best fitted by a log-normal curve. (D) Instantaneous speed distributions, resulting from all analyzed Myosin II null cells within micropillar array (red: *d*-runs, grey: *rm*-modes). The speed distribution of *rm*-modes is best fitted by a log-normal curve.

Micropillar array : interpillar distance		**<10µm**	**>15µm**
Number of data points in the statistics	n	4623	2189
Exponent of the l-MSD power-law			
all	$<\alpha>$	1.64	1.45
dir-runs	$<\alpha_d>$	1.82	1.81
rm-modes	$<\alpha_r>$	1.26	1.31
Instantaneous velocity (µm/s)			
all	$<V>$	0.040	0.040
dir-runs	$<V_d>$	0.059	0.048
rm-modes	$<V_r>$	0.031	0.035
***dir*-runs lifetime (s)**	$<\tau_d>$	120	108
***dir*-run length (µm)**	$l_{d,\,cut}$	8.4	5.8

Table S1: Parameters describing the cell motility within micropillar networks of varying interpillar distance, corresponding to Fig. 3, Fig. S1 to S3.

B. Supplementary Text

Text S1. Details concerning cell motility behavior on flat PDMS

We studied the spontaneous migration of single cells (N = 27 cells, n = 10968 data points) on flat PDMS substrates by applying both a global analysis and the time-resolved TRAnSpORT algorithm. The characteristic features are illustrated in Fig. 2 and summarized in Table 1.

Global MSD analysis The global MSD functions, calculated over the entire trajectories, yield information about the cell migration type as a whole. A typical example function is given in Fig. 2N (corresponding track shown in insert). The motion is superdiffusive at short time scales (power-law trend exponent $\alpha_1 = 1.70 >> 1$) and diffusive at long time scales (power-law trend exponent $\alpha_2 = 1.09 \approx 1$), with a typical cross-over time from one regime to the other of $\tau_C = 256s$. The statistics on all cells yield average values of the same order: $\langle \alpha_1 \rangle = 1.54 \pm 0.18$, $\langle \alpha_2 \rangle = 1.11 \pm 0.31$ and $\langle \tau_C \rangle = 207 \pm 124s$. But this global analysis can only give a rough representation of cell migration as a random walk which, at short time scales, involves periods of rather directed migration. While the double fitting of the global MSD functions can only yield one characteristic cross-over time τ_C, our high resolution analysis of local motion types enables to separate the contributions of different migration modes, and to analyze the distribution of their precise characteristics: local MSD exponent, velocity and lifetime.

Local MSD analysis The α-landscape of local MSD functions (Fig. 2B) shows a succession of ballistic-like (local exponent of the power law, $\alpha \geq 1.65$) and diffusive-like ($\alpha \leq 1.35$) phases, with an overall tendency towards superdiffusion. By correlating the values of the local exponent α and angle persistency $\Delta\phi$ (see Materials and Methods), we were able to split the cell migration trajectories into two specific motility types: *dir* -runs, or phases of directed motion, which also exhibit high velocities ($\langle v_d = 0.076 \mu m \cdot s^{-1} \rangle$), separated by *rm*-modes, i.e. slower phases of random probing ($\langle v_r = 0.057 \mu m \cdot s^{-1} \rangle$).

The distribution of the local exponents α confirms the tendency of cells to move in a directed way: it peaks around 1.8 (Fig. 2F). A second small bump in the distribution can be seen for α_1, which corresponds to the *rm*-modes, during which the cell randomly probes its environment and has a diffusive-like trajectory.

Moreover, the angle distributions of both phases, whether on a short or a long time scale, exhibit no particular trend in direction. So, the symmetry breaking, that triggers each directed phase, is isotropic.

Text S2. Details concerning cell motility behaviour within regular and homogeneous micropillar arrays

To extract the effect of micron-scale topography on amoeboid motility, we placed the cells in dense, regular arrays of micropillars (Fig. 1E). In this case, the cells have to squeeze in between pillars to move.

We studied the characteristics of the two motility modes, i.e. directed runs and random probing (N = 13 cells, n = 5178 data points), and noticed very clear differences to the spontaneous motion on a 2D surface, as illustrated in Fig. 2 and Table 1. The fact that the flat reference and the microstructured substrates consist of the same material (PDMS) ensures that any difference in cell motility can only be attributed to the well-defined micron-scale topography, and not to different chemical properties of the surface.

Global MSD analysis As seen for the typical cell track analyzed on Fig. 2O, the global MSD function exhibits a trend towards subdiffusion at long time scales ($\alpha_2 = 0.77$). Even though the motion remains superdiffusive at short time scales, the exponent is lower than on flat ($\alpha_1 = 1.23$), and the cross-over time is shorter ($\tau_C = 188s$). So the cell overall migration is less effcient, and results in a more confined random walk. This is confirmed by average values: $\langle \alpha_1 \rangle = 1.31 \pm 0.14$, $\langle \alpha_2 \rangle = 0.87 \pm 0.32$ and $\langle \tau_C \rangle = 176 \pm 84s$. Also, the average cell velocity is lower than the reference velocity on flat: $\langle V \rangle_{pil} = 0.043 \mu m \cdot s^{-1}$ instead of $\langle V \rangle_{ref} = 0.057 \mu m \cdot s^{-1}$ (Fig. 2C).

Local MSD analysis After splitting into the two motility modes, we observe that the velocity of the *dir*-runs is larger $\langle V \rangle_{pil} = 0.082 \mu m \cdot s^{-1}$ instead of $\langle V \rangle_{ref} = 0.077 \mu m \cdot s^{-1}$ on flat. The *dir*-runs are much less frequent: within this very dense micropillar environment, they represent only 10% of the total time, as opposed to 35% for the spontaneous motility on a flat substrate. This tendency of cells to move in a more erratic way is reflected in the α exponent distribution (Fig. 2G): it is peaked around 1.5, significantly lower than the reference distribution. The *dir*–run lifetimes decay much faster: fitting by a simple exponential yields a characteristic time of $\langle \tau_d \rangle = 111s$.

Finally, the most striking trend can be seen in the angle distribution of *dir*–runs (Fig. 2C): in such a topographical environment, the directed modes of motion do not occur in random directions any more. They are guided along the main axes of the lattice. The random probing modes, however, remain randomly oriented. This already shows that topographical features act as a guide for cell migration only by directing the ballistic phases of motion.

Here, our observation of cell migration in two drastically different topographical situations gives hints about the way motion is affected by dense microstructures, but does not yield information about the transition from one topographical environment to the other.

To study the migration transition at the interface between flat and microstructured environments, we recorded Dd cell migration within scarcely distributed islands of micropillar arrays, separated by large areas of flat surface. In addition to the dynamical motility features, we observed the steady-state spatial distribution of cells in environments where different topographical environments are adjoining, evidencing trends in cell localization towards microstructures.

Text S3. Details concerning the motility behavior within low-density islands of micropillars

We analyzed cell migration (N = 27 cells, n = 10220 data points) in environments composed of low-density islands of micropillars (Fig. 1I, 3C). Out of the 27 cells, 14 (50%) were observed to be in contact with pillars for all or a fraction of their trajectory. Out of them, 9 (75% of the contacting cells) displayed losing and re-establishing contact with pillars several times[1].

3.1. Description of typical motion features

Typical trajectories observed for cells migrating at the interface between the 2D surface and the 3D micropillars are presented in Fig. S3A, C and E (and the corresponding movies S3, S5, and S4 respectively).

The slow and probing cells display "stick-and-go". When a cell reaches a pillar, it typically stays in contact with it for more than 200 s (Fig. S3A,E). Upon detaching from it, the cell starts crawling on the flat PDMS, and performs a sequence of short *dir*-/*rm*-phases, going back and forth for 200 to 500 s (Fig. S3B).

Within that searching time, a "stick-and-go" motion feature was observed for most cells (60% of the contacting cells) as a typical reaction to pillar contact: either the cell finds another pillar, which becomes the new starting point of the next "stick-and-go" series, or the cell finds its way back to the same pillar. In both cases, long interaction times between cell and pillar are observed.

The revisiting events, often displayed several times successively, are a very interesting feature. Indeed, one would not expect them in the frame of an optimized searching process, and show that the complex internal signalling cascades triggered by prolonged micropillar contact (over 500 s) result in cell relocalization towards topographical features. The fact that cells have a certain biochemical "memory" of the obstacles they were in contact with can however be linked to another motion feature: on flat substrates, motion of the cell along its own previous track can sometimes be observed (Fig. 1DA and movie S1), reflecting the cell sensing of its own left-over protein traces [*SR1*].

The fast and efficiently migrating cells are deflected by pillars. If no other pillar is present in the cell vicinity after 500 to 1000 s of probing, the cell repolarizes in a new direction and starts a new *dir*–run (Fig. S3C, D: orange arrows). A cell which is then in a fast *dir*-run displays a very different reaction to pillar contact: when touching a pillar, it is deflected by it, but does not show any slowing down nor does it stay in contact with the micropillar for more than 100 s (Fig. S3C-D: white arrow, and movie S5). On the flat part of the substrate, such cells are exhibiting very frequent and long *dir*-runs, with a speed higher than $0.1\mu m \cdot s^{-1}$, which is at the upper limit of the overall speed distribution, since 70% of the *dir*-run speed values on 2D substrates are between zero and $\langle V_d \rangle + 3\sigma_{Vd} = 0.093\mu m \cdot s^{-1}$ (Fig. 2B).

3.2. Statistical results

Partition coefficient between flat and microstructured substrates Once given the possibility to stay on the flat part of the substrate or to come in contact with micropillars, the cells tend to spend a very high fraction of their time in contact with pillars. The overall statistics yield a "relocalization" factor R of almost 3 (Fig. 3D). This factor is calculated as the ratio of two

[1] We also observed cells staying on top of a pillar and turning around it on its edge, which proves that *D. discoideum* cells are sensitive to the high curvature present at the border between pillar top and pillar walls. Such a case can be observed in movies S3 and S9.

dimensionless numbers:

$$R = \frac{T_{cont}/T_{tot}}{\sum_{pil}/\sum_{flat}}$$

The numerator, T_{cont}/T_{tot}, represents the fraction of the time the cells spend in contact with pillars, for all measured cells: T_{cont} takes into account all the time points for which the contact index is greater than 1 and T_{tot} is the total number of time points recorded. The denominator, $\sum_{pil}/\sum_{flat}$ represents the fraction of the surface available for cells in the neighbourhood of pillars, as measured on the images. So the fact that we obtain a factor $R = 2.75$ means that the cells tend to stay almost three times longer in contact with pillars than on flat.

If we take into account the additional surface offered by pillar walls in the vertical direction, we retrieve a surface ratio $\sum_{pil}/\sum_{flat}$ twice larger. It results in a new ratio $R_{addsurf} \cong 1$. So, by taking into account the additional surface offered to the cells by the pillar walls, we obtain an equal distribution between cells on flat and cells in contact with microstructures.

This relocalization factor can also be seen as a steady-state partition coefficient: when considering the dynamical behavior of cells at the transition between a flat and a microstructured landscape, this coefficient represents the trend for cells to remain in one or the other topographical environment. We show here that the observed partition can be explained exclusively by an additional surface effect.

Cell attraction to micropillars as a dynamical effect The velocity distributions obtained from the time points during which the cells touch 1 or 2 pillars are slightly shifted towards smaller values, compared to the reference velocity distribution for cells migrating on flat PDMS (Fig. 3E). Indeed, the velocity distributions peak around $0.03\mu m \cdot s^{-1}$ for the flat case, $0.02\mu m \cdot s^{-1}$ for $I_{cont} = 1$ and $0.01\mu m \cdot s^{-1}$ for $I_{cont} = 2$. However, this shift is small compared to the standard deviation of the distributions (around $0.04\mu m \cdot s^{-1}$ in all three cases), and the median velocity stays within 10% around the value of $0.05\mu m \cdot s^{-1}$, as shown in insert of Fig. 3E.

In summary, the observed cell motion while in contact with a micropillar results from the superposition of the following counter-acting effects:

(1) an "additionally available surface" effect: the probability that random protrusions touch a surface is increased. If a pseudopod establishes contact with the micropillar surface, feedback loops are activated locally and trigger actin polymerization which stabilizes the pseudopod. The resulting probability that this pseudopod becomes a leading one, and hence the probability that a *dir*-run starts, is increased.

(2) a "three-dimensional motion" effect: when a *dir*-run starts upon micropillar contact, the cell moves partially in the *Z*-direction, but the *Z*-component of the velocity cannot be measured by our standard epifluorescence observations. Hence, the measured velocity is lower than the actual one and the proportion of *dir*-runs retrieved by analysis is also biased towards lower values, due to the undetected motion in the *Z*-direction.

The first effect tends to increase the proportion of directed motion modes, while the second tends to decrease it. As a result, a very slight decrease can be observed both in the velocity, as described above, and in the *dir*-run frequency upon pillar contact: from 35% on flat to 27% when touching two pillars (Fig. 3F).

This amoeboid search mechanism and its modification upon topography sensing appears strikingly similar to the one described for bacterial chemotaxis: the run-and-tumble mechanism and its modification upon gradient sensing have long been shown [*SR3, SR4*]. The underlying theory is that the spontaneous motion consists of a superposition of two states, each being a random walk on a different length scale (the directed runs result in long-distance bacteria motion while the tumble modes allow for local probing), and that chemotaxis results from a shift in the distribution of the state frequency. But in the case of bacteria, due to their small size, the long and directed runs remain randomly oriented, and the cell motion up-gradient only results from a

switch in frequency. While here, the orientation of the *dir*-runs is biased by topography sensing, because contact of one cell protrusion with a surface polarizes the cell into a new preferred direction.

3.4 Cell interaction with individual micropillars

Within the framework of our two-state model, we investigated interaction types of single cells with individual micropillars. As elaborated above for cell behavior on flat substrates, we observed the following characteristics:

(i) for cells which are initially in the random probing *rm*-mode: they frequently return to pillars after first pillar contact (Fig. S3A-B, S3C-D and movies S3 and S4). This revisiting effect can occur several times successively, which is not expected in the framework of an optimized search process. This proves that a cell in the probing *rm*-mode will be very sensitive to the positive feedback provided by contact with a pillar: the pillar walls offer additional surface in the third dimension (see scheme in Fig. 3B) and as a consequence, the probability for a randomly extruded cell protrusion to become the leading pseudopod is increased by pillar contact. The new direction of the cell motion is hence biased towards the pillar and the cell seems to be trapped around it.

(ii) for cells which are initially in a fast *dir*-run: upon touching a pillar, their *dir*-run is deflected. They slightly slow down and repolarize into a new direction after pillar contact. They rarely stay in contact with the micropillar for more than 100 s (Fig. S3E-F and movie S5) and do not remain trapped around it.

Micropillars as an indicator of the cell migration type

These two drastically different ways of reacting to micron-scale topography allow us to easily distinguish between two extreme types of motility states, while the cell-to-cell variability observed on a flat substrate is harder to categorize. Indeed, on flat substrates, the run-and-search motility includes all cell behaviors, ranging from the slow and probing cells to the fast and efficient cells. However, when in contact with a microstructure, cells in the *rm*-mode appear to be very sensitive to topographical cues and stay in contact with them for very long periods (over 500 s), tending to return when they do not find another pillar close-by. Conversely, the fast and efficient *dir*-run cells merely sense the topography and react to it with no other trajectory modification than a deflection.

Also, since cells can switch from one migration type to the other (as evidenced in Fig. S3C and movie S5), the presence of a topographical feature in the cell neighbourhood allows for the rapid discrimination between one or the other cell behavior type. Interestingly, the fact that a cell hits the pillar head-on or sideways is not decisive for the type of reaction triggered (see Movies S3 and S5: two cells hit a pillar sideways, but for one cell this results in a prolonged contact (S3), whereas it leads to deflection for the other cell (S5)).

C. Supplementary References

SR1. M. Schindl, et al., *Biophys*. J., 1995,**68**, 1177.
SR2. T. Bretschneider, et al., *Biophys. J.*, 2009, **96**, 2888.
SR3. H. Berg and D. Brown, *Nature*, 1972, **239**, 500.
SR4. J. Adler, *Science,* 1969, **166**, 1588 .

D. Description of the supporting movies

Movie S1 *D. discoideum* cell migrating freely on a flat PDMS substrate.
Corresponding trajectory : Fig. 1D.

Movie S2 *D. discoideum* cell migrating within a dense rectangular array of PDMS micropollars. In order to be easily followed over the movie, the cell is highlighted by a white circle.
Corresponding trajectory: Fig. 1E.

Movie S3 *D. discoideum* cell migrating at the frontier between micropillars and the flat substrate. The cell displays revisiting "stick-and-go", by coming back twice to the same pillar.
Corresponding trajectory: Fig. S3A.

Movie S4 *D. discoideum* cell migrating at the frontier between an island of micropillars and the flat substrate. The cell displays "stick-and-go", by going from pillar to pillar.
Corresponding trajectory: Fig. S3C.

Movie S5 *D. discoideum* cell migrating at the frontier between an island of micropillars and the flat substrate. The cell switches from slow (pillar-sensitive) motility at the beginning to fast and efficient (pillar-insensitive) migration at the end (where pillar contact only results in the deflection of the cell). The switching time at which the cell repolarizes and starts migrating fast into a new direction is indicated by a red circle.
Corresponding trajectory: Fig. S3E.
(N.B. : movie has been rotated by 90 deg. with regard to the original images, for file compressing reasons)

Movie S6 *D. discoideum* cell migrating within a micropillar array of low density, going from pillar to pillar. The cell of interest is highlighted by a white circle at the beginning of the movie.
Corresponding trajectory: Fig. S1A.

Movie S7 *D. discoideum* cell migrating within a micropillar array of low density, going from pillar to pillar and covering a very long distance.
Corresponding trajectory: Fig. S1B.
(N.B. : movie has been rotated by 90 deg. with regard to the original images, for file compressing reasons)

Movie S8 *D. discoideum* cells migrating within a micropillar array of intermediate density, at which they can touch two pillars at the same time. The cells then display stalling behavior, highlighted by magenta circles.
Corresponding trajectory: Fig. S2D.

C Protocols

C.1 Preparation of microstructured six-channel slides

Before you start

- PLL-g-PEG(2) ($1mg/ml$) in 10mM Hepes buffer with 150mM NaCl @ pH 7.4
- Fibronectin ($50\mu g/ml$) in PBS
- Ibidi sticky slide $VI^{0.4}$
- Ibidi foils uncoated: Place them on top of a channel slide to mark the positions of the channels, make a little scratch into the protective foil using tweezers to facilitate later lift-off
- PDMS stamp of the wafer

Plasma treatment

- Place the cut PDMS stamps onto the foil.
- Settings: Timer 3.0, Power 5.5, oxygen plasma
- Carefully push the stamps onto the foil after plasma treatment (use tweezers).

PLL-g-PEG (work under the flow-hood from now on)

- You will need very small amounts of PEG (about $1.5\mu l$ per stamp), $8\mu l$ are sufficient for PEGylation of a whole channel slide.
- Place the pipette at the top edge of the slightly tilted stamp-on-foil. The PEG will flow under the stamp by capillary action.
- Incubate for 30 minutes at room-temperature.

Combine sticky foil and channel slide.

- Place the channel slide upside down, remove foil so that the adhesive film is exposed.
- Remove the PDMS stamp from the foil (use tweezers).
- To hit the target, try to guide the foil with your fingers when putting the foil onto the channel slide.

- Carefully rub the foil so that it sticks perfectly to the channel slide, using your fingers or tweezers. Do not destroy the protective film!
- Rinse the channels with deionized water.
- Fill the channels with fibronectin ($50\mu l$ per channel). Incubate for 1h at room-temperature.
- Rinse three times with deionized water.

You can store the channels (either empty or filled with PBS) in the fridge for up to two weeks (or more).
A cell number of 10000 per channel was optimal for A549 and Huh7 cells.
Seed the cells at least three hours prior to transfection, pipette cell solution through the channel several times. The total volume per channel is $150\mu l$.
Add anti-evaporation oil before long-term measurements.

C.2 Preparation of microchannels and GFP calibration

The height of the PDMS channels should not exceed the focal depth of the microscope objective! Focal depth of Nikon 10x objective is about $8.5\mu m$, PDMS wafers No.3 ($3000rpm, 25s, h = 7.9\mu m$) and No.5 ($5000rpm, 10s, h = 6.6\mu m$) can be used

- Cut PDMS (1:10/crosslinker:monomer/w:w) stamp and punch holes at the end of the channels
- Plasma cleaner: PDMS stamp (upside down) and cover slip, next to each other, oxygen plasma, power 4.0, timer 2.0
- Immediately turn the stamp and carefully press it onto the cover slip.
- Add $3\mu l$ F108 (3mM) into one of the two reservoirs of each channel. After some time, add $3\mu l$ to the opposite reservoirs.
- Put a stripe of scotch tape on top of the channel slide and incubate for $20h$ @ $4°$.
- Rinse each channel three times with PBS, dry them under nitrogen flow.
- Fill the five channels with a concentration serial dilution of GFP stock solution, cover channel slide with scotch tape or PDMS monomer.
- Use the same microscope settings as in the corresponding experiments (exposure time, fluorescence lamp intensity, binning, intensity range)!
- Capture background for correction.
- Run z-stacks (about $30\mu m$ range). The middle one of the eight consecutively brightest slices ($1\mu m$ step size) supposedly is the middle of the channel.

Determine the concentration of the GFP stock solution using either fluorescence correlation spectroscopy or using the Nanodrop spectrophotometer.
Correct for autofluorescence and uneven illumination using the background correction of SingleCellTracker and determine mean intensities per pixel of equally sized regions of interest using ImageJ. For the 10-fold Nikon objective, one voxel is roughly $3.17\mu m^3$. Assign grey values per pixel to numbers of GFP molecules (see figure 2.6), thus generating a calibration curve for the conversion of grey values to numbers of GFP molecules (see also supplementary of [54]).

C.3 A very general transfection protocol

Before you start:
OptiMEM transfection medium and *Lipofectamine 2000* ® (Lipofectamine) should be at room-temperature, vortex Lipofectamine shortly before usage.
Thaw and vortex DNA.
mRNA should be kept on ice, use RNAse-free pipette-tips!

- Prepare equal volumes of nucleic acid and lipid in transfection medium in the two steps to follow:
- Dilute nucleic acid in transfection medium.
- Dilute Lipofectamine in transfection medium (for most experiments, $2.5\mu l$ of Lipofectamine were used for each μg of nucleic acid. However, this ratio is cell-line dependent and may have to be optimised according to the purpose of the experiment.)
- Wait for five minutes.
- Thoroughly mix both solutions in equal amounts, wait for 20 minutes for lipoplex formation.
- Wash cells with PBS before transfection, add OptiMEM before adding the lipoplex solution.
- Carefully add the lipoplexes to the cells, change medium to growth medium or microscopy medium after incubation at 37°.

C.4 Cell preparation for flow cytometry measurements

C.4.1 Preparation for immediate measurements

The amounts are given on a per-well-basis for a standard 24-well plate. Prepare ice-cold PBS for the resuspension step.

- Remove medium.
- Rinse with PBS.
- Add $100\mu l$ Trypsin/EDTA, incubate at 37° for 3 minutes.
- Using a $100\mu l$ pipette, carefully detach the cells.
- Add $1ml$ medium to stop the Trypsin/EDTA.
- Centrifuge for 5 minutes at $400g$ and 21°.
- Resuspend the cells in $900\mu l$ ice-cold PBS containing 10% FCS.

Store the samples on ice and measure immediately.

C.4.2 Paraformaldehyde fixation

The amounts are given on a per-well-basis for a standard six-well plate. Prepare ice-cold PBS for the resuspension step. Wear protection and work under the extraction hood for all paraformaldehyde-steps!

- Remove medium.
- Rinse with PBS.
- Add $300\mu l$ Trypsin/EDTA, incubate at 37° for 3 minutes.
- Using a $100\mu l$ pipette, carefully detach the cells.
- Add $1ml$ medium to stop the Trypsin/EDTA.
- Centrifuge for 5 minutes at $400g$ and 21°.
- Remove supernatant.
- Resuspend the cell pellet in $1ml$ ice-cold PBS.
- Under the hood: mix with $1ml$ 4% paraformaldehyde.

Samples prepared this way can be stored at 4° for more than two weeks. However, because this protocol is based on toxic paraformaldehyde, the alternative preparation described in the preceding protocol should be preferred whenever possible.

Bibliography

[1] O. Shimomura, F. H. Johnson, and Y. Saiga. Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, aequorea. *Journal of Cellular and Comparative Physiology*, 59(3):223–239, 1962.

[2] M. Chalfie, Tu Y., G. Euskirchen, W. W. Ward, and D. C. Prasher. Green fluorescent protein as a marker for gene expression. *Science*, 263:802–805, 1994.

[3] R. Y. Tsien. The green fluorescent protein. *Annual Review of Biochemistry*, 67(1):509–544, 1998.

[4] G. H. Patterson and J. Lippincott-Schwartz. A photoactivatable GFP for selective photolabeling of proteins and cells. *Science*, 297(5588):1873–1877, 2002.

[5] C. Okita, M. Sato, and T. Schroeder. Generation of optimized yellow and red fluorescent proteins with distinct subcellular localization. *BioTechniques*, 36(3):418–22, 424, 2004.

[6] J. S. Paige, K. Y. Wu, and S. R. Jaffrey. RNA mimics of green fluorescent protein. *Science*, 333(6042):642–646, 2011.

[7] S. W. Hell and J. Wichmann. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Optics Letters*, 19(11):780–782, 1994.

[8] E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess. Imaging intracellular fluorescent proteins at nanometer resolution. *Science*, 313(5793):1642–1645, 2006.

[9] S. T. Hess, T. P. K. Girirajan, and M. D. Mason. Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophysical Journal*, 91(11):4258–4272, 2006.

[10] M. J. Rust, M. Bates, and X. Zhuang. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature Methods*, 3(10):793–796, 2006.

[11] L. S. Barak and W. W. Webb. Fluorescent low density lipoprotein for observation of dynamics of individual receptor complexes on cultured human fibroblasts. *The Journal of Cell Biology*, 90(3):595–604, 1981.

[12] A. M. Femino, F. S. Fay, K. Fogarty, and R. H. Singer. Visualization of single RNA transcripts in situ. *Science*, 280(5363):585–590, 1998.

[13] Y. B. Ya'ara Ben-Ari, N. Kinor, A. Mor, T. Tsukamoto, D. L. Spector, R. H. Singer, and Y. Shav-Tal. The life of an mRNA in space and time. *Journal of Cell Science*, 123(10):1761, 2010.

[14] N. Ruthardt, D. C. Lamb, and C. Bräuchle. Single-particle tracking as a quantitative microscopy-based approach to unravel cell entry mechanisms of viruses and pharmaceutical nanoparticles. *Molecular Therapy*, 19(7):1199–1211, 2011.

[15] I. Golding, J. Paulsson, S. M. Zawilski, and E. C. Cox. Real-time kinetics of gene activity in individual bacteria. *Cell*, 123(6):1025–1036, 2005.

[16] L. Cai, N. Friedman, and X. S. Xie. Stochastic protein expression in individual cells at the single molecule level. *Nature*, 440(7082):358–362, 2006.

[17] J. Yu, J. Xiao, X. Ren, K. Lao, and X. S. Xie. Probing gene expression in live cells, one protein molecule at a time. *Science*, 311(5767):1600–1603, 2006.

[18] G.-W. Li and X. S. Xie. Central dogma at the single-molecule level in living cells. *Nature*, 475(7356):308–315, 2011.

[19] A. P. Cotrim and B. J. Baum. Gene therapy: Some history, applications, problems, and prospects. *Toxicologic Pathology*, 36(1):97–103, 2008.

[20] M. L. Edelstein, M. R. Abedi, J. Wixon, and R. M. Edelstein. Gene therapy clinical trials worldwide 1989-2004 - an overview. *The Journal of Gene Medicine*, 6(6):597–602, 2004.

[21] K. Strebhardt and A. Ullrich. Paul ehrlich's magic bullet concept: 100 years of progress. *Nature Reviews Cancer*, 8(6):473–480, 2008.

[22] M.-C. Hung, L. Huang, and E. Wagner. *Nonviral vectors for gene therapy.* Academic Press, 1999.

[23] D. Schaffert and E. Wagner. Gene therapy progress and prospects: synthetic polymer-based systems. *Gene Therapy*, 15(16):1131–1138, 2008.

[24] N. Kamaly, Z. Xiao, P. M. Valencia, A. F. Radovic-Moreno, and O. C. Farokhzad. Targeted polymeric therapeutic nanoparticles: design, development and clinical translation. *Chemical Society Reviews*, 41(7):2971–3010, 2012.

[25] M. R. Capecchi. High efficiency transformation by direct microinjection of DNA into cultured mammalian cells. *Cell*, 22(2, Part 2):479–488, 1980.

[26] E. Neumann, M. Schaefer-Ridder, Y. Wang, and P. H. Hofschneider. Gene transfer into mouse lyoma cells by electroporation in high electric fields. *The EMBO Journal*, 1(7):841, 1982.

[27] P. L. Felgner, T. R. Gadek, M. Holm, R. Roman, H. W. Chan, M. Wenz, J. P. Northrop, G. M. Ringold, and M. Danielsen. Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure. *Proceedings of the National Academy of Sciences*, 84(21):7413–7417, 1987.

[28] M. X. Tang, C. T. Redemann, and F. C. Szoka. In vitro gene delivery by degraded polyamidoamine dendrimers. *Bioconjugate Chemistry*, 7(6):703–714, 1996.

[29] M. Jordan, A. Schallhorn, and F. M. Wurm. Transfecting mammalian cells: optimization of critical parameters affecting calcium-phosphate precipitate formation. *Nucleic Acids Research*, 24(4):596–601, 1996.

[30] N. S. Templeton, D. D. Lasic, P. M. Frederik, H. H. Strey, D. D. Roberts, and G. N. Pavlakis. Improved DNA: liposome complexes for increased systemic delivery and gene expression. *Nature Biotechnology*, 15(7):647–652, 1997.

[31] W.-C. Tseng, F. R. Haselton, and T. D. Giorgio. Transfection by cationic liposomes using simultaneous single cell measurements of plasmid delivery and transgene expression. *Journal of Biological Chemistry*, 272(41):25641–25647, 1997.

[32] M. Lindgren, M. Hällbrink, A. Prochiantz, and Ü. Langel. Cell-penetrating peptides. *Trends in Pharmacological Sciences*, 21(3):99–103, 2000.

[33] F. Scherer, M. Anton, U. Schillinger, J. Henke, C. Bergemann, A. Krüger, B. Gänsbacher, and C. Plank. Magnetofection: enhancing and targeting gene delivery by magnetic force in vitro and in vivo. *Gene Therapy*, 9(2):102, 2002.

[34] D. W. Pack, A. S. Hoffman, S. Pun, and P. S. Stayton. Design and development of polymers for gene delivery. *Nature Reviews Drug Discovery*, 4(7):581–593, 2005.

[35] C. Dufès, I. F. Uchegbu, and A. G. Schätzlein. Dendrimers in gene delivery. *Advanced Drug Delivery Reviews*, 57(15):2177–2202, 2005.

[36] J. Dobson. Gene therapy progress and prospects: magnetic nanoparticle-based gene delivery. *Gene Therapy*, 13(4):283–287, 2006.

[37] K. de Bruin, N. Ruthardt, K. von Gersdorff, R. Bausinger, E. Wagner, M. Ogris, and C. Bräuchle. Cellular dynamics of EGF receptor-targeted synthetic viruses. *Molecular Therapy*, 15(7):1297–1305, 2007.

[38] H. Debus, P. Baumhof, J. Probst, and T. Kissel. Delivery of messenger RNA using poly(ethylene imine)-poly(ethylene glycol)-copolymer blends for polyplex formation: Biophysical characterization and in vitro transfection properties. *Journal of Controlled Release*, 148(3):334–343, 2010.

[39] V. J. Schüller, S. Heidegger, N. Sandholzer, P. C. Nickels, N. A. Suhartha, S. Endres, C. Bourquin, and T. Liedl. Cellular immunostimulation by CpG-sequence-coated DNA origami structures. *ACS Nano*, 5(12):9696–9702, 2011.

[40] D. Smith, V. Schüller, C. Engst, J. Rädler, and T. Liedl. Nucleic acid nanostructures for biomedical applications. *Nanomedicine*, 8(1):105–121, 2012.

[41] W. Li and F. C. Szoka Jr. Lipid-based nanoparticles for nucleic acid delivery. *Pharmaceutical Research*, 24(3):438–449, 2007.

[42] A. J. Hamilton and D. C. Baulcombe. A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science*, 286(5441):950–952, 1999.

[43] S. M. Elbashir, J. Harborth, W. Lendeckel, A. Yalcin, K. Weber, and T. Tuschl. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature*, 411(6836):494–498, 2001.

[44] R. W. Malone, P. L. Felgner, and I. M. Verma. Cationic liposome-mediated RNA transfection. *Proceedings of the American Thoracic Society*, 86(16):6077–81, 1989.

[45] B. A. Sullenger and E. Gilboa. Emerging clinical applications of RNA. *Nature*, 418(6894):252–258, 2002.

[46] A. Yamamoto, M. Kormann, J. Rosenecker, and C. Rudolph. Current prospects for mRNA gene delivery. *European Journal of Pharmaceutics and Biopharmaceutics*, 71(3):484–489, 2009.

[47] J. Rejman, G. Tavernier, N. Bavarsad, J. Demeester, and S. C. De Smedt. mRNA transfection of cervical carcinoma and mesenchymal stem cells mediated by cationic carriers. *Journal of Controlled Release*, 147(3):385–391, 2010.

[48] G. Tavernier, O. Andries, J. Demeester, N. N. Sanders, S. C. De Smedt, and J. Rejman. mRNA as gene therapeutic: How to control protein expression. *Journal of Controlled Release*, 150(3):238–247, 2011.

[49] M. S. D. Kormann, G. Hasenpusch, M. K. Aneja, G. Nica, A. W. Flemmer, S. Herber-Jonat, M. Huppmann, L. E. Mays, M. Illenyi, A. Schams, M. Griese, I. Bittmann, R. Handgretinger, D. Hartl, J. Rosenecker, and C. Rudolph. Expression of therapeutic proteins after delivery of chemically modified mRNA in mice. *Nature Biotechnology*, 29(2):154–157, 2011.

[50] M. Esteller. Non-coding RNAs in human disease. *Nature Reviews Genetics*, 12(12):861–874, 2011.

[51] M. Mockey, C. Gonçalves, F. P. Dupuy, F. M. Lemoine, C. Pichon, and P. Midoux. mRNA transfection of dendritic cells: Synergistic effect of ARCA mRNA capping with poly(A) chains

in cis and in trans for a high protein expression level. *Biochemical and Biophysical Research Communications*, 340(4):1062–1068, 2006.

[52] S. Zou, K. Scarfo, M. H. Nantz, and J. G. Hecker. Lipid-mediated delivery of RNA is more efficient than delivery of DNA in non-dividing cells. *International Journal of Pharmaceutics*, 389(1):232–243, 2010.

[53] S. McLenachan, D. Zhang, A. B. A. Palomo, M. J. Edel, and F. K. Chen. mRNA transfection of mouse and human neural stem cell cultures. *PloS One*, 8(12):e83596, 2013.

[54] C. Leonhardt, G. Schwake, T. R. Stögbauer, S. Rappl, J.-T. Kuhr, T. S. Ligon, and J. O. Rädler. Single-cell mRNA transfection studies: Delivery, kinetics and statistics by numbers. *Nanomedicine: Nanotechnology, Biology and Medicine*, 10(4):679–688, 2014.

[55] S. S. Diebold, T. Kaisho, H. Hemmi, S. Akira, and C. R. e Sousa. Innate antiviral responses by means of TLR7-mediated recognition of single-stranded RNA. *Science*, 303(5663):1529–1531, 2004.

[56] V. Hornung, J. Ellegast, S. Kim, K. Brzózka, A. Jung, H. Kato, H. Poeck, S. Akira, K.-K. Conzelmann, and M. Schlee. 5'-triphosphate RNA is the ligand for RIG-I. *Science*, 314(5801):994–997, 2006.

[57] A. Pichlmair, O. Schulz, C. P. Tan, T. I. Näslund, P. Liljeström, F. Weber, and C. R. e Sousa. RIG-I-mediated antiviral responses to single-stranded RNA bearing 5'-phosphates. *Science*, 314(5801):997–1001, 2006.

[58] S. Holtkamp, S. Kreiter, A. Selmi, P. Simon, M. Koslowski, C. Huber, O. Tureci, and U. Sahin. Modification of antigen-encoding RNA increases stability, translational efficacy, and T-cell stimulatory capacity of dendritic cells. *Blood*, 108(13):4009–17, 2006.

[59] A. N. Kuhn, M. Diken, S. Kreiter, B. Vallazza, Ö. Türeci, and U. Sahin. Determinants of intracellular RNA pharmacokinetics. *RNA Biology*, 8(1):35–43, 2011.

[60] I. Strobel, S. Berchtold, A. Götze, U. Schulze, G. Schuler, and A. Steinkasserer. Human dendritic cells transfected with either RNA or DNA encoding influenza matrix protein M1 differ in their ability to stimulate cytotoxic T lymphocytes. *Gene Therapy*, 7(23), 2000.

[61] D. Boczkowski, J. Lee, S. Pruitt, and S. Nair. Dendritic cells engineered to secrete anti-GITR antibodies are effective adjuvants to dendritic cell-based immunotherapy. *Cancer Gene Therapy*, 16(12):900–911, 2009.

[62] K. Takahashi and S. Yamanaka. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*, 126(4):663–676, 2006.

[63] K. Takahashi, K. Tanabe, M. Ohnuki, M. Narita, T. Ichisaka, K. Tomoda, and S. Yamanaka. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*, 131(5):861–872, 2007.

[64] L. Warren, P. D. Manos, T. Ahfeldt, Y.-H. Loh, H. Li, F. Lau, W. Ebina, P. K. Mandal, Z. D. Smith, and A. Meissner. Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. *Cell Stem Cell*, 7(5):618–630, 2010.

[65] L. Warren, Y. Ni, J. Wang, and X. Guo. Feeder-free derivation of human induced pluripotent stem cells with messenger RNA. *Science Reports*, 2, 2012.

[66] P. K. Mandal and D. J. Rossi. Reprogramming human fibroblasts to pluripotency using modified mRNA. *Nature Protocols*, 8(3):568–582, 2013.

[67] N. L. Garneau, J. Wilusz, and C. J. Wilusz. The highways and byways of mRNA decay. *Nature Reviews Molecular Cell Biology*, 8(2):113–126, 2007.

[68] Y. Gong, A. O. Ogunniyi, and J. C. Love. Massively parallel detection of gene expression in single cells using subnanolitre wells. *Lab on a Chip*, 10(18):2334–2337, 2010.

[69] M. Bengtsson, A. Ståhlberg, P. Rorsman, and M. Kubista. Gene expression profiling in single cells from the pancreatic islets of langerhans reveals lognormal distribution of mRNA levels. *Genome Research*, 15(10):1388–1392, 2005.

[70] J. Gertsch, M. Güttinger, O. Sticher, and J. Heilmann. Relative quantification of mRNA levels in jurkat T cells with RT-real time-PCR (RT-rt-PCR): New possibilities for the screening of anti-inflammatory and cytotoxic compounds. *Pharmaceutical Research*, 19(8):1236–1243, 2002.

[71] T. Nolan, R. E. Hands, and S. A. Bustin. Quantification of mRNA using real-time RT-PCR. *Nature Protocols*, 1(3):1559–1582, 2006.

[72] R. A. Gutiérrez, R. M. Ewing, J. M. Cherry, and P. J. Green. Identification of unstable transcripts in arabidopsis by cDNA microarray analysis: rapid decay is associated with a group of touch-and specific clock-controlled genes. *Proceedings of the National Academy of Sciences*, 99(17):11513–11518, 2002.

[73] M. Kenzelmann, S. Maertens, M. Hergenhahn, S. Kueffer, A. Hotz-Wagenblatt, L Li, S. Wang, C. Ittrich, T. Lemberger, and R. Arribas. Microarray analysis of newly synthesized RNA in cells and animals. *Proceedings of the National Academy of Sciences*, 104(15):6164–6169, 2007.

[74] Z. Dölken, L.and Ruzsics, B. Rädle, C. C. Friedel, R. Zimmer, J. Mages, R. Hoffmann, P. Dickinson, T. Forster, and P. Ghazal. High-resolution gene expression profiling for simultaneous kinetic parameter analysis of RNA synthesis and decay. *RNA*, 14(9):1959–1972, 2008.

[75] B. Schwanhäusser, D. Busse, N. Li, G. Dittmar, J. Schuchhardt, J. Wolf, W. Chen, and M. Selbach. Global quantification of mammalian gene expression control. *Nature*, 473(7347):337–342, 2011.

[76] C. Miller, B. Schwalb, K. Maier, D. Schulz, S. Dümcke, B. Zacher, A. Mayer, J. Sydow, L. Marcinowski, and L. Dölken. Dynamic transcriptome analysis measures rates of mRNA synthesis and decay in yeast. *Molecular Systems Biology*, 7(1), 2011.

[77] I. K. Dimov, R. Lu, E. P. Lee, J. Seita, D. Sahoo, S. Park, I. L. Weissman, and L. P. Lee. Discriminating cellular heterogeneity using microwell-based RNA cytometry. *Nature Communications*, 5, 2014.

[78] D. B. Wheeler, A. E. Carpenter, and D. M. Sabatini. Cell microarrays and RNA interference chip away at gene function. *Nature Genetics*, 37:S25–S30, 2005.

[79] A. Azioune, M. Storch, M. Bornens, M. Thery, and M. Piel. Simple and rapid process for single cell micro-patterning. *Lab on a Chip*, 9(11):1640–1642, 2009.

[80] E. Nagoshi, C. Saini, C. Bauer, T. Laroche, F. Naef, and U. Schibler. Circadian gene expression in individual fibroblasts: cell-autonomous and self-sustained oscillators pass time to daughter cells. *Cell*, 119(5):693–705, 2004.

[81] N. Rosenfeld, J. W. Young, U. Alon, P. S. Swain, and M. B. Elowitz. Gene regulation at the single-cell level. *Science*, 307(5717):1962–1965, 2005.

[82] A. Sigal, R. Milo, A. Cohen, N. Geva-Zatorsky, Y. Klein, Y. Liron, N. Rosenfeld, T. Danon, N. Perzov, and U. Alon. Variability and memory of protein levels in human cells. *Nature*, 444(7119):643–646, 2006.

[83] B. Neumann, M. Held, U. Liebel, H. Erfle, P. Rogers, R. Pepperkok, and J. Ellenberg. High-throughput RNAi screening by time-lapse imaging of live human cells. *Nature Methods*, 3(5):385–390, 2006.

[84] M. Halter, A. Tona, K. Bhadriraju, A. L. Plant, and J. T. Elliott. Automated live cell imaging of green fluorescent protein degradation in individual fibroblasts. *Cytometry Part A*, 71A(10):827–834, 2007.

[85] J. A. Megerle, G. Fritz, U. Gerland, K. Jung, and J. O. Rädler. Timing and dynamics of single cell gene expression in the arabinose utilization system. *Biophysical Journal*, 95(4):2103–15, 2008.

[86] J. C. Locke and M. B. Elowitz. Using movies to analyse gene circuit dynamics in single cells. *Nature Reviews Microbiology*, 7(5):383–92, 2009.

[87] D. Muzzey and A. van Oudenaarden. Quantitative time-lapse fluorescence microscopy in single cells. *Annual Review of Cell and Developmental Biology*, 25(1):301–327, 2009.

[88] G. Schwake, S. Youssef, J. T. Kuhr, S. Gude, M. P. David, E. Mendoza, E. Frey, and J. O. Rädler. Predictive modeling of non-viral gene transfer. *Biotechnology and Bioengineering*, 105(4):805–13, 2009.

[89] T. Schroeder. Long-term single-cell imaging of mammalian stem cells. *Nature Methods*, 8(4s):S30–S35, 2011.

[90] P. J. Röttgermann, A. Piera-Alberola, and J. O. Rädler. Cellular self-organization on micro-structured surfaces. *Soft Matter*, 10(14):2397–2404, 2014.

[91] D. D. Carlo and L. P. Lee. Dynamic single-cell analysis for quantitative biology. *Analytical Chemistry*, 78(23):7918–7925, 2006.

[92] S. L. Spencer, S. Gaudet, J. G. Albeck, J. M. Burke, and P. K. Sorger. Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis. *Nature*, 459(7245):428–432, 2009.

[93] A. Raj and A. van Oudenaarden. Nature, nurture, or chance: Stochastic gene expression and its consequences. *Cell*, 135(2):216–226, 2008.

[94] J. M. Pedraza and A. van Oudenaarden. Noise propagation in gene networks. *Science*, 307(5717):1965–1969, 2005.

[95] M. Tigges, T. T. Marquez-Lago, J. Stelling, and M. Fussenegger. A tunable synthetic mammalian oscillator. *Nature*, 457(7227):309–312, 2009.

[96] M. Leisner, L. Bleris, J. Lohmueller, Z. Xie, and Y. Benenson. Rationally designed logic integration of regulatory signals in mammalian cells. *Nature Nanotechnology*, 5(9):666–670, 2010.

[97] T. S. Ledley and F. D. Ledley. Multicompartment, numerical model of cellular events in the pharmacokinetics of gene therapies. *Human Gene Therapy*, 5(6):679–691, 1994.

[98] D. S. Friend, D. Papahadjopoulos, and R. J. Debs. Endocytosis and intracellular processing accompanying transfection mediated by cationic liposomes. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1278(1):41–50, 1996.

[99] C. M. Varga, K. Hong, and D. A. Lauffenburger. Quantitative analysis of synthetic gene delivery vector design properties. *Molecular Therapy*, 4(5):438–446, 2001.

[100] C. M. Varga, N. C. Tedford, M. Thomas, A. M. Klibanov, L. G. Griffith, and D. A. Lauffenburger. Quantitative comparison of polyethylenimine formulations and adenoviral vectors in terms of intracellular gene delivery processes. *Gene Therapy*, 12(13):1023–1032, 2005.

[101] H. Harashima, Y. Shinohara, and H. Kiwada. Intracellular control of gene trafficking using liposomes as drug carriers. *European Journal of Pharmaceutical Sciences*, 13(1):85–89, 2001.

[102] H. Kamiya, H. Akita, and H. Harashima. Pharmacokinetic and pharmacodynamic considerations in gene therapy. *Drug Discovery Today*, 8(21):990–996, 2003.

[103] G. A. Banks, R. J. Roselli, R. Chen, and T. D. Giorgio. A model for the analysis of nonviral gene therapy. *Gene Therapy*, 10(20):1766–1775, 2003.

[104] A. T. Dinh, C. Pangarkar, T. Theofanous, and S. Mitragotri. Understanding intracellular transport processes pertinent to synthetic gene delivery via stochastic simulations and sensitivity analyses. *Biophysical Journal*, 92(3):831–46, 2007.

[105] A. Mahmutovic, D. Fange, O. G. Berg, and J. Elf. Lost in presumption: stochastic reactions in spatial models. *Nature Methods*, 9(12):1163–1166, 2012.

[106] J. Fisher and T. A. Henzinger. Executable cell biology. *Nature Biotechnology*, 25(11):1239–1249, 2007.

[107] H. D. Summers, P. Rees, M. D. Holton, M. Rowan Brown, S. C. Chappell, P. J. Smith, and R. J. Errington. Statistical analysis of nanoparticle dosing in a dynamic cellular system. *Nature Nanotechnology*, 6(3):170–174, 2011.

[108] E. M Ozbudak, M. Thattai, I. Kurtser, A. D. Grossman, and A. van Oudenaarden. Regulation of noise in the expression of a single gene. *Nature Genetics*, 31(1):69–73, 2002.

[109] X. Darzacq, Y. Shav-Tal, V. de Turris, Y. Brody, S. M. Shenoy, R. D. Phair, and R. H. Singer. In vivo dynamics of RNA polymerase II transcription. *Nature Structural & Molecular Biology*, 14(9):796–806, 2007.

[110] W. J. Blake, M. Kaern, C. R. Cantor, and J. J. Collins. Noise in eukaryotic gene expression. *Nature*, 422(6932):633–637, 2003.

[111] V. Shahrezaei and P. S. Swain. Analytical distributions for stochastic gene expression. *Proceedings of the National Academy of Sciences*, 105(45):17256–17261, 2008.

[112] G. Pesole, G. Grillo, A. Larizza, and S. Liuni. The untranslated regions of eukaryotic mRNAs: Structure, function, evolution and bioinformatic tools for their analysis. *Briefings in Bioinformatics*, 1(3):236–249, 2000.

[113] S. N. Cohen, A. C. Y. Chang, H. W. Boyer, and R. B. Helling. Construction of biologically functional bacterial plasmids in vitro. *Proceedings of the National Academy of Sciences*, 70(11):3240–3244, 1973.

[114] S. P. Goff and P. Berg. Construction of hybrid viruses containing SV40 and lambda phage DNA segments and their propagation in cultured monkey cells. *Cell*, 9(4, Part 2):695–705, 1976.

[115] C. E. Thomas, A. Ehrhardt, and M. A. Kay. Progress and problems with the use of viral vectors for gene therapy. *Nature Reviews Genetics*, 4(5):346–358, 2003.

[116] M. C. Pedroso de Lima, S. Simões, P. Pires, H. Faneca, and N. Düzgünes. Cationic lipid-DNA complexes in gene delivery: from biophysics to biological applications. *Advanced Drug Delivery Reviews*, 47(2-3):277–294, 2001.

[117] B. Dalby, S. Cates, A. Harris, E. C. Ohki, M. L. Tilkins, P. J. Price, and V. C. Ciccarone. Advanced transfection with lipofectamine 2000 reagent: primary neurons, siRNA, and high-throughput applications. *Methods*, 33(2):95–103, 2004.

[118] M. C. Woodle and D. D. Lasic. Sterically stabilized liposomes. *Biochimica et Biophysica Acta (BBA)-Reviews on Biomembranes*, 1113(2):171–199, 1992.

[119] A. Gabizon, R. Catane, B. Uziely, B. Kaufman, T. Safra, R. Cohen, F. Martin, A. Huang, and Y. Barenholz. Prolonged circulation time and enhanced accumulation in malignant exudates of doxorubicin encapsulated in polyethylene-glycol coated liposomes. *Cancer Research*, 54(4):987–992, 1994.

[120] V. P. Torchilin. Recent advances with liposomes as pharmaceutical carriers. *Nature Reviews Drug Discovery*, 4(2):145–160, 2005.

[121] S. Rudorf and J. O. Rädler. Self-assembly of stable monomolecular nucleic acid lipid particles with a size of 30 nm. *Journal of the American Chemical Society*, 134(28):11652–8, 2012.

[122] J. O. Rädler, I. Koltover, T. Salditt, and C. R. Safinya. Structure of DNA-cationic liposome complexes: DNA intercalation in multilamellar membranes in distinct interhelical packing regimes. *Science*, 275(5301):810–814, 1997.

[123] C. R. Safinya. Structures of lipid-DNA complexes: supramolecular assembly and gene delivery. *Current Opinion in Structural Biology*, 11(4):440–448, 2001.

[124] K. A. Mislick and J. D. Baldeschwieler. Evidence for the role of proteoglycans in cation-mediated gene transfer. *Proceedings of the National Academy of Sciences*, 93(22):12349–12354, 1996.

[125] L. C. Mounkes, W. Zhong, G. Cipres-Palacin, T. D. Heath, and R. J. Debs. Proteoglycans mediate cationic liposome-DNA complex-based gene delivery in vitro and in vivo. *Journal of Biological Chemistry*, 273(40):26164–26170, 1998.

[126] D. Vanhecke and M. Janitz. High-throughput gene silencing using cell arrays. *Oncogene*, 23(51):8353–8358, 2004.

[127] A. Piera Alberola and J. O. Rädler. The defined presentation of nanoparticles to cells and their surface controlled uptake. *Biomaterials*, 30(22):3766–3770, 2009.

[128] I. A. Khalil, K. Kogure, H. Akita, and H. Harashima. Uptake pathways and subsequent intracellular trafficking in nonviral gene delivery. *Pharmacological Reviews*, 58(1):32–45, 2006.

[129] W.-C. Tseng, F. R. Haselton, and T. D. Giorgio. Mitosis enhances transgene expression of plasmid delivered by cationic liposomes. *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression*, 1445(1):53–64, 1999.

[130] D. Magde, E. Elson, and W. W. Webb. Thermodynamic fluctuations in a reacting system - measurement by fluorescence correlation spectroscopy. *Physical Review Letters*, 29(11):705–708, 1972.

[131] E. P. Petrov and P. Schwille. *State of the Art and Novel Trends in Fluorescence Correlation Spectroscopy*, volume 6 of *Springer Series on Fluorescence*, book section 32, pages 145–197. Springer Berlin Heidelberg, 2008.

[132] D. Davenport and J. A. C. Nicol. Luminescence in hydromedusae. *Proceedings of the Royal Society of London. Series B - Biological Sciences*, 144(916):399–411, 1955.

[133] G. H. Patterson, S. M. Knobel, W. D. Sharif, S. R. Kain, and D. W. Piston. Use of the green fluorescent protein and its mutants in quantitative fluorescence microscopy. *Biophysical Journal*, 73(5):2782–2790, 1997.

[134] A. Revzin, R. G. Tompkins, and M. Toner. Surface engineering with poly(ethylene glycol) photolithography to create high-density cell arrays on glass. *Langmuir*, 19(23):9855–9862, 2003.

[135] M. R. Dusseiller, D. Schlaepfer, M. Koch, R. Kroschewski, and M. Textor. An inverted microcontact printing method on topographically structured polystyrene chips for arrayed micro-3-D culturing of single cells. *Biomaterials*, 26(29):5917–5925, 2005.

[136] J. R. Rettig and A. Folch. Large-scale single-cell trapping and imaging using microwell arrays. *Analytical Chemistry*, 77(17):5628–5634, 2005.

[137] A. Khademhosseini, R. Langer, J. Borenstein, and J. P. Vacanti. Microscale technologies for tissue engineering and biology. *Proceedings of the National Academy of Sciences*, 103(8):2480–2487, 2006.

[138] H. Kim, R. E. Cohen, P.T. Hammond, and D.J. Irvine. Live lymphocyte arryas for biosensing. *Advanced Functional Materials*, 16(10):1313–1323, 2006.

[139] M. Ochsner, M. R. Dusseiller, H. M. Grandin, S. Luna-Morris, M. Textor, V. Vogel, and M. L. Smith. Micro-well arrays for 3D shape control and high resolution analysis of single cells. *Lab on a Chip*, 7(8):1074–1077, 2007.

[140] K. Franke, I. Kurth, M. Bornhauser, C. Werner, and T. Pompe. Biomimetic microcavities based on poly(dimethylsiloxane) elastomers. *Soft Matter*, 5(18):3505–3510, 2009.

[141] A.-K. Marel, S. Rappl, A. Piera-Alberola, and J. O. Rädler. Arraying cell cultures using PEG-DMA micromolding in standard culture dishes. *Macromolecular Bioscience*, 13(5):595–602, 2013.

[142] D. Falconnet, A. Koenig, F. Assi, and M. Textor. A combined photolithographic and molecular-assembly approach to produce functional micropatterns for applications in the biosciences. *Advanced Functional Materials*, 14(8):749–756, 2004.

[143] J. Fink, M. Thery, A. Azioune, R. Dupont, F. Chatelain, M. Bornens, and M. Piel. Comparative study and improvement of current cell micro-patterning techniques. *Lab on a Chip*, 7(6):672–680, 2007.

[144] S. Cookson, N. Ostroff, W. L. Pang, D. Volfson, and J. Hasty. Monitoring dynamics of single-cell gene expression over multiple cell cycles. *Molecular Systems Biology*, 1, 2005.

[145] A. Paganin-Gioanni, E. Bellard, J. M. Escoffre, M. P. Rols, J. Teissié, and M. Golzio. Direct visualization at the single-cell level of siRNA electrotransfer into cancer cells. *Proceedings of the National Academy of Sciences*, 108(26):10443–10447, 2011.

[146] G. G. Stokes. On the change of refrangibility of light. *Philosophical Transactions of the Royal Society of London*, 142:463–562, 1852.

[147] N. C. Shaner, P. A. Steinbach, and R. Y. Tsien. A guide to choosing fluorescent proteins. *Nature Methods*, 2(12):905–909, 2005.

[148] S. Youssef, S. Gude, and J. O. Rädler. Automated tracking in live-cell time-lapse movies. *Integrative Biology*, 3(11):1095 – 1101, 2011.

[149] Z. Jericevic, B. Wiese, J. Bryan, and L. C. Smith. Validation of an imaging system: Steps to evaluate and validate a microscope imaging. *Methods in Cell Biology*, 30:47, 1989.

[150] M. A. Model and K. E. Healy. Quantification of the surface density of a fluorescent label with the optical microscope. *Journal of Biomedical Materials Research*, 50(1):90–96, 2000.

[151] M. Schwarzfischer, C. Marr, J. Krumsiek, P. S. Hoppe, T. Schroeder, and F. J. Theis. *Efficient fluorescence image normalization for time lapse movies.* 2011.

[152] M. J. Fulwyler. Electronic separation of biological cells by volume. *Science*, 150(3698):910–911, 1965.

[153] O. Andries, M. De Filette, J. Rejman, S. C. De Smedt, J. Demeester, M. Van Poucke, L. Peelman, C. Peleman, T. Lahoutte, and N. N. Sanders. Comparison of the gene transfer efficiency of mRNA/GL67 and pDNA/GL67 complexes in respiratory cells. *Molecular Pharmaceutics*, 9(8):2136–2145, 2012.

[154] C. Barreau, S. Dutertre, L. Paillard, and H. B. Osborne. Liposome-mediated RNA transfection should be used with caution. *RNA*, 12(10):1790–1793, 2006.

[155] J. Zabner, A. J. Fasbender, T. Moninger, K. A. Poellinger, and M. J. Welsh. Cellular and molecular barriers to gene transfer by a cationic lipid. *Journal of Biological Chemistry*, 270(32):18997–19007, 1995.

[156] D.T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, 81(25):2340–2361, 1977.

[157] S. McLenachan, J. P. Sarsero, and P. A. Ioannou. Flow-cytometric analysis of mouse embryonic stem cell lipofection using small and large DNA constructs. *Genomics*, 89(6):708–720, 2007.

[158] A. Schloßbauer, A. M. Sauer, V. Cauda, A. Schmidt, H. Engelke, U. Rothbauer, K. Zolghadr, H. Leonhardt, C. Bräuchle, and T. Bein. Cascaded photoinduced drug delivery to cells from multifunctional core-shell mesoporous silica. *Advanced Healthcare Materials*, 1(3):316–320, 2012.

[159] G. Seelig, D. Soloveichik, D. Y. Zhang, and E. Winfree. Enzyme-free nucleic acid logic circuits. *Science*, 314(5805):1585–8, 2006.

[160] M. Zuker and P. Stiegler. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. *Nucleic Acids Research*, 9(1):133–148, 1981.

[161] J. S. McCaskill. The equilibrium partition function and base pair binding probabilities for RNA secondary structure. *Biopolymers*, 29(6 7):1105–1119, 1990.

[162] I. L. Hofacker, W. Fontana, P. F. Stadler, L. S. Bonhoeffer, M. Tacker, and P. Schuster. Fast folding and comparison of RNA secondary structures. *Monatshefte für Chemie/Chemical Monthly*, 125(2):167–188, 1994.

[163] D. Arcizet, S. Capito, M. Gorelashvili, C. Leonhardt, S. Youssef, S. Rappl, and D. Heinrich. Contact-controlled amoeboid motility induces dynamic cell trapping in 3D-microstructured surfaces. *Soft Matter*, 8(5):1473–1481, 2012.

Full List of Publications

- **Carolin Leonhardt***, Gerlinde Schwake*, Tobias R. Stögbauer, Susanne Rappl, Jan-Timm Kuhr, Thomas S. Ligon, and Joachim O. Rädler. Single-cell mRNA transfection studies: Delivery, kinetics and statistics by numbers. *Nanomedicine: Nanotechnology, Biology, and Medicine*, 10, 679-688 (2014). *equal contribution

- Mehrije Ferizi*, **Carolin Leonhardt***, Christian Meggle, Manish K. Aneja, Carsten Rudolph, Christian Plank, and Joachim O. Rädler. Stability Analysis of chemically modified mRNA using micropattern-based Single-Cell Arrays. *to be submitted*. *equal contribution

- Thomas S. Ligon, **Carolin Leonhardt**, and Joachim O. Rädler. Multi-Level Kinetic Model of mRNA Delivery via Transfection of Lipoplexes. *PLoS ONE*, 9(9): e107148 (2014).

- Delphine Arcizet*, Sofia Capito*, Mari Gorelashvili, **Carolin Leonhardt**, Marion Vollmer, Simon Youssef, Susanne Rappl, and Doris Heinrich. Contact-controlled amoeboid motility induces dynamic cell trapping in 3D-microstructured surfaces. *Soft Matter*, 8, 1473-1481 (2012). *equal contribution

Danksagung

Mein Dank gilt vielen, von denen ich hier hoffentlich niemanden vergessen werde. Meinem Doktorvater **Prof. Joachim Rädler** für die Betreuung und Unterstützung während der gesamten Zeit meiner Promotion und seine menschliche Art, Chef zu sein. **Dr. Jan-Timm Kuhr** sowie **Mehrije Ferizi, Dr. Manish Aneja, Prof. Carsten Rudolph, Dr. Johannes Geiger, Prof. Christian Plank** und **Dr. Jan Hasenauer** für die produktive Zusammenarbeit. **Marilena Pinto** vom IDK. Allen aktuellen und ehemaligen **"Rädlers"**, die das so angenehme Arbeitsklima prägen und geprägt haben. Besonderer Dank gilt dabei **Gerlinde Schwake** für den guten Zuspruch, die angenehme Abnahme der ein oder anderen Laborsündenbeichte, für die Sause und alles. Hierfür danke ich auch **Dr. Tobias Stögbauer**. **Max Albert** und **Dr. Susanne Kempter** für Hilfe im Labor bzw. Langlaufunterricht. **Dr. D. H.** "for being soft on the issue and so tough on the person", was sehr lehrreich war. **Dr. Delphine Arcizet, Felix Keber** und **Dr. Marcus Otten**, die immer ein offenes Ohr und Rat für mich hatten. **Dr. Thomas Ligon** für die gute Zusammenarbeit und die ein oder andere interessante Anekdote. **Bässem Salem, Stephan Rudolph** und **Laura Lechtenberg**, von denen ich als Betreuerin lernen durfte. **Christian Meggle** für sein klickarmes Bildauswertungsprogramm. Den **Rennrädlers** sowie den **aktiven Pausiererinnen**, hierunter vor allem **Sonja Westermeyer**, denn Sport tut gut. **Dr. Alicia Piera-Alberola** für die Antwort auf viele Fragen in einer so ruhigen Art. **Margarete Meixner**, **Andreas Mader**, **Peter Roettgermann** und **Tobias Preiß** für ihre große Hilfsbereitschaft. Der **Augenweide** fürs Weiden vor meinen Augen und dem **Herrn Kommissar** nicht nur dafür, dass er mir den Oktopus gezeigt hat. **Dr. David Smith** und **Philipp Nickels** før danske spas und die Teepausen. **Dr. Silvia Milani** per parlare italiano con me. **Angelika Kardinal**, **Thomas Nicolaus**, **Dr. Martin Benoit** und **Dr. Stefan Scheuer** vom Lehrstuhl Gaub. Meinen CSVS Freundinnen **Susanne Rappl, Dr. Verena Schüller** und **Svenja Lippok**, die viel mit mir gelacht, manchmal aber auch mit mir geweint haben, die mich hin und wieder ertragen haben und ohne die die letzten wie die kommenden Jahre ganz anders aussähen. Ihr gehört auch in den nächsten Satz.
Zuguterletzt, eigentlich aber vor allem will ich mich auch bei all jenen Menschen außerhalb der Uni, die für mich wichtig sind, bedanken - dafür, dass sie mich in der Zeit dieser Arbeit begleitet haben und mich dabei immer wieder daran erinnert haben, was mir wirklich wichtig ist.