

---

# Dynamics of stochastic membrane rupture events: effects on drug delivery and gene expression

Maria Pamela Dobay (nee David)

---



München 2012



---

# Dynamics of stochastic membrane rupture events: effects on drug delivery and gene expression

Maria Pamela Dobay (nee David)

---

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

vorgelegt von  
Maria Pamela Dobay (nee David)  
aus Quezon City, Philippinen

München, den 26. April 2012

Erstgutachter: Prof. Dr. Joachim O. Rädler

Zweitgutachter: Prof. Dr. Ulrich Gerland

Tag der mündlichen Prüfung: 2. Juli 2012

# Contents

|  |            |
|--|------------|
| <b>Zusammenfassung</b>   | <b>xi</b>  |
| <b>Summary</b>   | <b>xiv</b> |
| <b>1 Introduction</b>  | <b>1</b>   |
| 1.1 Motivation . . . . .   | 1          |
| 1.2 State-of-the art . . . . .   | 1          |
| 1.3 Overview of the thesis . . . . .   | 3          |
| <b>2 Modeling non-viral gene transfer using a hybrid stochastic model</b>                                    | <b>5</b>   |
| 2.1 Single cell expression data-derived distributions of exogenous gene expression characteristics . . . . . | 6          |
| 2.2 Use of a hybrid deterministic/stochastic model to describe delivery processes                            | 7          |
| 2.3 Limitations of differential equation models . . . . .  | 8          |
| <b>3 Modeling stochastic cellular processes: dealing with compartments and complex networks</b>              | <b>13</b>  |
| 3.1 Introduction to pi calculus . . . . .  | 14         |
| 3.2 The Stochastic Pi Machine (SPiM): a pi calculus implementation . . . . .                                 | 15         |
| 3.2.1 Introduction to SPiM . . . . .   | 15         |
| 3.2.2 Rule-based modeling in SPiM . . . . .  | 16         |
| 3.2.3 The Gillespie algorithm . . . . .  | 17         |
| 3.3 Brane calculi: modeling stochastic membrane processes . . . . .  | 20         |
| 3.4 A Projective Brane calculus with activate, bud and mate as primitive actions (PABM) . . . . .            | 22         |
| <b>4 Modeling gold NP uptake and intracellular distribution</b>  | <b>29</b>  |
| 4.1 Choice of gold nanoparticles as a case study . . . . .   | 30         |
| 4.2 Nanoparticle aggregation, uptake, and combinatorial explosion . . . . .                                  | 30         |
| 4.3 Pi calculus model of gold NP uptake and intracellular distribution . . . . .                             | 31         |
| <b>5 Experiments and modeling of endosomal escape events</b>   | <b>33</b>  |
| 5.1 Characterization of endosomal release of a lipid-coated nanoparticle reporter system . . . . .           | 34         |

---

|          |   |            |
|----------|---|------------|
| 5.1.1    | Generation of lipid-coated nanoparticles . . . . .  | 34         |
| 5.1.2    | Reporter system characterization and experiments . . . . .  | 34         |
| 5.2      | Stochastic pi calculus model of PpIX-induced endosomal escape . . . . .   | 36         |
| <b>6</b> | <b>Design considerations for nanoparticles with protein or peptide modifications for targeted delivery and controlled release</b> | <b>43</b>  |
| 6.1      | Modeling minimal requirements for HA-mediated endosomal release . . . . .   | 44         |
| 6.2      | Outlook . . . . .   | 45         |
| 6.2.1    | NP uptake insights . . . . .  | 45         |
| 6.2.2    | Future work in nanoparticle design . . . . .  | 46         |
| <b>A</b> | <b>Original publication P1</b>  | <b>49</b>  |
| <b>B</b> | <b>Original publication P2</b>  | <b>61</b>  |
| <b>C</b> | <b>Original publication P3</b>  | <b>87</b>  |
| <b>D</b> | <b>Original publication P4</b>  | <b>101</b> |
| <b>E</b> | <b>Original publication P5</b>  | <b>111</b> |
| <b>F</b> | <b>Manuscript M1</b>  | <b>123</b> |
|          | <b>Acknowledgements</b>   | <b>155</b> |

# List of Figures

|     |  |    |
|-----|--|----|
| 2.1 | Exogenous gene expression in single cells . . . . .                    | 7  |
| 2.2 | GFP expression statistics . . . . .                                    | 10 |
| 2.3 | Transfection efficiency for PEI and lipofectamine . . . . .            | 11 |
| 3.1 | Rule based modeling . . . . .  | 16 |
| 3.2 | Brane calculus reduction rules . . . . .                               | 25 |
| 3.3 | Differences between Brane and Projective Brane Calculi . . . . .       | 26 |
| 3.4 | PABM domains . . . . .   | 27 |
| 3.5 | PABM reduction rules . . . . .   | 27 |
| 4.1 | Spatio-temporal distribution of gold NPs . . . . .                     | 32 |
| 5.1 | CMS-based reporter system . . . . .                                    | 37 |
| 5.2 | Time distributions of PpIX-mediated endosomal escape . . . . .         | 38 |
| 5.3 | Endosome size distribution and NP colocalization intensities . . . . . | 39 |
| 5.4 | PpIX model flow . . . . .  | 40 |
| 5.5 | SPiM model fits and parameter sensitivity analysis . . . . .           | 41 |
| 6.1 | Cell-type specific uptake II: time-course experiments . . . . .        | 46 |
| 6.2 | Tracker and carrier NP Design . . . . .                                | 47 |



# List of Tables

|     |  |    |
|-----|--|----|
| 3.1 | Objects and operations of pi calculus . . . . .    | 14 |
| 3.2 | Biochemical and pi objects . . . . .               | 15 |
| 3.3 | Objects and operations of Brane calculus . . . . . | 21 |
| 3.4 | Reduction rules of Brane calculus . . . . .        | 22 |
| 3.5 | Objects and operations of PABM . . . . .           | 23 |
| 3.6 | Reduction rules of PABM . . . . .                  | 24 |
| 5.1 | Gauss fit parameters of NP uptake . . . . .        | 36 |



# Zusammenfassung

Die Wechselwirkung von Membranen mit kleinsten Masseteilchen ist ein zentrales Element in der Pharmakotherapie wie auch in der Interaktion von Viren und Nanopartikeln mit Zellen. Sie ist gekennzeichnet durch topologische Übergänge wie zum Beispiel die Endozytose, den Austritt aus dem Endosom sowie den Eintritt in den Zellkern. Bisherige Rechenmethoden eigneten sich allerdings nicht für eine umfassende Beschreibung solcher topologischen Reaktionen. Die vorliegende Arbeit beschreibt und vergleicht die Anwendung mathematischer und regelbasierter Modellierungsmethoden (Pi-Kalkül) auf Daten über die Aufnahme und Verteilung von Nanopartikeln, um das Verhalten nanopartikelförmiger Systeme und der zugehörigen zellulären Einsatzorte zu verstehen. Zudem wurde eine formale Modellierungssprache zur Beschreibung topologischer Übergänge entwickelt. Die von uns aufgestellten Modelle wurden mit Hilfe von Experimenten an Modellmembransystemen auf ihre Gültigkeit hin überprüft.

Im zweiten Kapitel wird ein erstes Beispiel für ein nicht-virales Gentransfersystem vorgestellt, in dem der intrazelluläre Transport von Teilchen zu einer direkt messbaren exogenen Genexpression führt. Hierzu wurde in Zusammenarbeit mit J.-T. Kuhr aus der Arbeitsgruppe von Prof. Dr. Erwin Frey ein Hybridmodell aus Differentialgleichungen und einem stochastischen Modell geschaffen. Dieses Modell wurde dazu verwendet, die Ursache für Schwankungen von Genexpressionsparametern wie der maximalen Anzahl grün fluoreszierender Proteine (GFP) pro Zelle, der Expressionsrate und des Zeitpunkts des Expressionsbeginns zu verstehen. Es erfasst nicht nur die Verteilung der Daten, sondern kann auch dazu verwendet werden, die Anzahl der tatsächlich transkribierten Plasmide abzuschätzen. Hierbei zeigt sich, dass diese Zahl deutlich kleiner ist als die Anzahl der ursprünglich durch die Zelle aufgenommenen Plasmide. Der größte Verlust an Vektoren vor dem Eintritt in den Zellkern kann dabei der unzureichenden Freisetzung aus den Endosomen zugeschrieben werden.

Die Vielfalt an kombinatorischen Möglichkeiten, die sich bei einer schrittweisen Beschreibung von Partikelaggregation bis zur Interaktion der Aggregate mit den verschiedenen Zellkompartimenten ergibt, limitiert die Anwendbarkeit von Differentialgleichungen. Aus diesem Grund wurde ein regelbasiertes Modell entwickelt, das die Komplexität des Problems auf die Anzahl der einzelnen Reaktionsmuster reduziert. Im dritten Kapitel wird aufgezeigt, inwiefern sich das stochastische Pi-Kalkül, welches eine formale und regelbasierte Sprache zur Modellierung von nebenläufigen, kommunizierenden Systemen ist, zur Modellierung biologischer Systeme eignet. Die Semantik des Pi-Kalküls bietet zwar einen

rudimentären Lösungsansatz für die Darstellung von Zellkompartimenten, jedoch keine optimale Beschreibung von komplexen Systemen mit tief verschachtelten Reaktionsräumen. Aus diesem Grund führen wir zusätzlich einen Formalismus, den sogenannten “projective brane calculus based on activate, bud and mate as primitive actions (PABM)”, ein. Dieser Formalismus berücksichtigt bestimmte Regeln bei membrantopologischen Umformungen sowie die Heterogenität (“patchiness”) der Membran.

Im vierten Kapitel verwenden wir diese Pi-Kalkül Modelle, um die intrazelluläre Verteilung von Goldnanopartikeln zu beschreiben. Die Lokalisierung solcher Nanopartikel in Zellkompartimenten kann mit Hilfe von Transmissionselektronenmikroskopie (TEM) gut visualisiert werden. Des Weiteren sind aufgrund von ausführlichen Charakterisierungsstudien über Goldnanopartikel kinetische Daten zu deren Aggregation, Aufnahme, endosomalem Austritt sowie Eintritt in den Zellkern entweder bereits vorhanden oder aber leicht abzuschätzen. Unsere Ergebnisse und Voraussagen demonstrieren, dass sich das hier verwendete Kalkül für die Modellierung relativ einfacher Probleme aus dem Bereich der Interaktion von Zellen mit Partikeln eignet. Im Gegensatz dazu wäre die Modellierung solcher Wechselwirkungen mit Hilfe von Differentialgleichungen bereits relativ komplex.

Im Anschluss hieran wurde ein Experiment modelliert und durchgeführt, in dem der topologische Übergang des Austritts aus dem Endosom kontrolliert werden kann. Die dazugehörigen Ergebnisse sind im fünften Kapitel zusammengefasst: In Zusammenarbeit mit A. Schmidt aus der Gruppe von Prof. Dr. Thomas Bein wurden lipidbeschichtete, kolloidal-, mesoporöse Nanopartikel, welche mit Protoporphyrin (PPIX) sowie Atto633-QSY21 konjugiert sind, verwendet, um zeitliche Verteilungen einzelner Austrittsereignisse aus Endosomen zu generieren. Die kontrollierte, lichtinduzierte Auflösung von Endosomen über die Aktivierung von PPIX lieferte überraschenderweise unterschiedliche Profile der Lysis in Fibroblasten und Karzinomzellen. Außerdem variierten die Verteilungen der Aufnahmezeiten dieser Nanopartikel. Um den Einfluss jener Faktoren auf Seiten der Zellen wie auch der Nanopartikel, welche die PPIX-induzierte Lysis beeinflussen, abschätzen zu können, wurde das Goldnanopartikelmodell erweitert. Details wie der unterschiedliche PPIX-Inhalt pro Nanopartikel, die Anzahl von Nanopartikeln je Endosom sowie die Nanopartikelaggregation wurden nun zusätzlich berücksichtigt. Unsere Ergebnisse zeigen, dass die zellabhängigen Verteilungen des Austritts aus Endosomen hauptsächlich durch die minimale Nettoanzahl von PPIX, die für die Auflösung eines Endosoms erforderlich ist, bestimmt werden. Dies steht im Einklang mit unseren Daten sowie Voraussagen aus einem weiteren theoretischen Modell über die Abhängigkeit des Zeitpunkts des endosomalen Austritts von der Größe des Endosoms.

In Kapitel 6 wird schließlich ein biologisches Membransystem modelliert, das die meisten Eigenschaften eines idealen Gentransfersystems besitzt: das Virus. Insbesondere modellierten wir die Dynamik des Beginns des Influenza Hämagglutinin(HA)-vermittelten endosomalen Austritts. In diesem Fall ist der Auslöser des topologischen Übergangs keine einfache Rate wie im Fall der lichtinduzierten Lysis, sondern die präzise Aggregation und Konformationsänderung von Membranproteinen. Wir verwenden hier ein PABM Modell, welches mit dem Modellchecker PRISM simuliert wird und die Unterschritte beim Aufbau der HA-Fusionspore beschreibt. Die besten Ergebnisse erhält man mit mindestens sechs

HA Trimeren pro Fusionspore, von denen drei eine Konformationsänderung durchlaufen müssen, um fusionsfördernd zu werden.

Zusammenfassend lässt sich sagen, dass die vorliegende Arbeit zeigt, wie man unter Verwendung regelbasierter Formalismen Modelle erstellen kann, die topologische Übergänge explizit beinhalten. Die mit solchen Modellen generierten Daten sowie die Voraussagen, welche durch sie getroffen werden können, sind von großem Wert für die Verbesserung künstlicher Gentransfersysteme.



# Summary

The interaction of membrane systems with particulate material, a recurring theme in drug delivery, and in virus- and nanoparticle (NP)-cell interaction, is characterized by topological transitions such as endocytosis, endosomal escape, and nuclear entry. However, current computational methods do not provide a way for handling topological reactions. In this work, we describe and compare the application of mathematical and rule-based, pi calculus models to uptake and delivery data in order to understand the behavior of nanoparticulate material with their corresponding cellular targets. We also developed a formal modeling language for specifically handling topological transformations. We subsequently designed and performed experiments on model membrane systems in order to test the validity of the created models.

In Chapter 2, we show a first example of artificial gene delivery, where the intracellular transfer of particulate material has a directly measurable outcome in the form of extrogenous gene expression. We created hybrid ordinary differential equation (ODE)/two-step stochastic model, in collaboration with J.-T. Kuhr the group of Prof. Dr. Erwin Frey, to understand the source of variation in the distributions of gene expression parameters, such as the maximum GFP per cell, expression rate and expression onset time. The model not only captures the data distributions, but can also be used to estimate the number of transcribed plasmids, which was found to be significantly less than the number of plasmids originally delivered to the cell. Of the processes upstream of nuclear entry, majority of the vector loss has been attributed to poor endosomal escape.

The application of ODEs, however, is limited by describing these upstream events, where a step-by-step description of how particles agglomerate and how these interact with the species in different cell compartments leads to combinatorial explosion. For this reason, we created rule-based models, which reduce complexity to the number of *distinct* reaction patterns. In Chapter 3, we show how stochastic pi calculus, a rule-based formal language for modeling concurrent mobile, communicating systems, can be used for modeling biological systems. Pi calculus semantics, while providing a rudimentary solution for representing compartments, is still not the best suited for manipulating complex systems with deeply-nested compartments. For this reason, we also introduced a formalism, a projective brane calculus based on activate, bud and mate as primitive actions (PABM), that has specific rules for handling membrane topological transformations, and that takes membrane patchiness into account.

In Chapter 4, we applied this pi calculus approach to model the intracellular distri-

bution of gold NPs. We chose gold NPs because these can be imaged using transmission electron microscopy (TEM) to obtain localization data across cellular compartments. Additionally, kinetic data for aggregation, uptake, endosomal escape, and nuclear entry are either available or easily estimated for gold NPs, owing to its extensive characterization. Data fits and predictions of NP intracellular distribution at time points used in independent gold NP studies demonstrate the adequacy of the calculus for modeling relatively simple problems of cell-particulate material interactions, which would have otherwise been already fairly difficult to model with ODEs.

We then created, and subsequently modeled, an experimental system where a topological transition, endosomal escape, is triggered on demand, results of which are summarized in Chapter 5. Time distributions of individual endosomal escape events were generated using a protoporphyrin (PPIX)- and Atto633-QSY21-conjugated, lipid-coated colloidal mesoporous silica (CMS) NP system, in a joint project with A. Schmidt from the group of Prof. Dr. Thomas Bein. Controlled, light-induced endosomal lysis through PPIX activation surprisingly yielded different lysis profiles in fibroblast and carcinoma cells, and also exhibited varying uptake profiles. In order to estimate the contributions of cell- and NP-related factors that affect PPIX-induced lysis, we extended the previous gold NP model to include details such as varying PPIX load per NP, NP load per endosome, and NP aggregation. Our results indicate that these cell-dependent endosomal escape profiles are primarily influenced by the minimum *net* number of PPIX required to lyse an endosome. This is consistent with our data, as well as predictions from an independent, theoretical model of escape time dependence on endosome size variance.

In Chapter 6, we finally model a biological membrane system that exhibits most of the characteristics the ideal gene delivery system: the virus. In particular, we model the early fusion dynamics of influenza hemagglutinin (HA)-mediated endosomal escape. In this case, the trigger of topological change is no longer governed by a simple rate, as the case was in light-induced rupture, but by the concerted aggregation and conformational change of membrane proteins. Here, we use a PABM model, simulated using the model checker PRISM, to create a model that includes individual steps of HA pore unit assembly. The best fit obtained was for a minimum of six HA trimers that form the fusion pore, of which three need to undergo a conformational change to become fusogenic. Collectively, this work shows how models that incorporate topological transitions explicitly can be created using rule-based formalisms; the type of data that can be obtained from such models; and the value of the predictions in terms of how these could be used in improving vector design.

# Chapter 1

## Introduction

### 1.1 Motivation

The ability to predict and control uptake and intracellular interactions of nanoparticulate material, which include artificial delivery vectors, nanoparticles (NPs), and viruses themselves, is a central theme in gene and drug delivery. To some extent, control is imbued by adding targeting sequences to artificial vectors and NPs, including cell- and organelle-specific ligands. The numerous steps and interactants in delivery, however, could easily result in unexpected vector behavior [1, 2]. Key to understanding these processes is the creation of quantitative, mechanistic models and experimental model systems, that allow us to dissect and interpolate between the steps in these processes, explain why certain behaviors arise, and eventually use these information in vector design. In this thesis, we focus on the construction of such models and model systems, as well as on the development of a modeling language that is specifically appropriate for detailing effects of topological transformations in the cell.

### 1.2 State-of-the art

Vector design currently entails an extensive evaluation of physico-chemical properties, as well as biological barriers that may affect vector performance [3]. Computer-aided vector design has been mostly used for predicting the so-called ADME parameters (adsorption, distribution, metabolism and excretion). On the design side, the quantitative structure-activity relationship (QSAR) approach, a multivariate regression model that attempts to link the molecular structure of compounds to its biological activity, is currently being adapted to NP activity and toxicity assessment. In particular, a focus is given on evaluating which existing data are appropriate for “nano-QSARs”, and which physico-chemical properties, or *descriptors*, have to be specifically obtained for NPs [4]. Previously used in applications ranging from drug discovery to chemical risk management, it is currently being adapted to identify key physical parameters of an NP, including surface parameters, geometry, size and composition, that can be used as predictors of its performance and

cytotoxicity [5]. Currently, nano-QSAR studies account for most of the modeling done in NP-related research. It has been used to predict the uptake [6] and toxicity [5, 7] of different types of metal oxide particles based on twelve structural descriptors, including area and volume parameters [6, 5], as well as detachment enthalpies of metal and gaseous cations from the NP [5]; QSARs have also been adapted to specifically predict the behavior of compounds with the lipid bilayer, and has been used in identifying factors that may influence carbon nanotube toxicity [8]. However, it is speculated that there will be no generalized, one-nano-QSAR-fits-all solution due to the extremely high variability of molecular structure, which presumably also translates to a similarly wide variety of action mechanisms; instead, it is surmised that different QSARs will be applied to different classes of NPs [4].

Statistical analysis has also been used to characterize the delivery of both protein-targeted and unmodified NPs; the uptake events in both cases were characterized by a broadly-spread Poisson distribution whose variance is greater than its mean. Such a distribution is consistent with heterogenous adsorption and internalization, i.e. uptake events are random [9]. The fates of NPs following cell division were also demonstrated to be both random and asymmetric; the asymmetry might be due to endosome association with microtubules [10, 9], in turn thought to arise from an evolved, protective mechanism of asymmetric partitioning of damaged proteins [11, 12, 9]. The implications of the study on dosing and delivery are very important: given that number of NPs *per endosome* is random, and that intrinsic cell-to-cell variability provides an additional layer of randomness to the delivery process, maximizing the number of *dose units* per cell is crucial – in this case, defined by the endosome; with the number of endosomes, which effectively limits the number of delivery trials, at around  $10^2$  per cell, the minimum dose variation that can be achieved is approximately 10%. Direct cytosolic delivery, however, has been reported to result in the internalization of up to  $10^6$  molecules, resulting in a variation of around 0.1% [13, 9]. Consequently, the latter mode of delivery is more desirable, if precise dosing is required [9].

Statistical models by themselves, however, do not the explain the mechanism behind a predicted outcome. This is where mechanical models, frequently in the form of ordinary differential equation (ODE) models, have been used. For instance, a combination of QSAR and molecular dynamics simulation was used for elucidating the most likely structural and dynamic contributors to carbon nanotube cellular toxicity [8]. Several ODE and thermodynamic models have also been constructed to understand how NP design, particularly, NP size and ligand density, affect cellular uptake [14, 15, 16], which have predicted optimal size and ligand density ranges for NPs. There have also been ODE models that detail viral [17] or delivery vector [18, 19] intracellular movement and interaction with cellular components, which have led to the identification of critical components or steps governing the delivery process. The fact that a number of biological events are driven by molecules with low copy numbers, stochastic representations have also been used to supplement continuum approaches [20]. Among the recent stochastic models relevant to our work describes protease-dependent endosomal escape of nonenveloped viruses, where the escape probability was shown to be dependent on the net number of proteases in the endosome and the

endosome size [21].

It is conceded, however, that more realistic and better-integrated approaches that could bridge the different scales of the problem – from atom to organism, are needed [22, 23]. The complexity of biological networks, however, is itself hampering the development of models that capture this “whole picture”. While ODE models have become a standard for monitoring fluxes of biological and biochemical systems, these suffer severely from problems of scale [20]. This is where the advantages of rule-based modeling, which uses reaction patterns, rather than individual reaction definitions, become evident. Rule based models prevent the combinatorial explosion that would otherwise occur with ODEs, and allow models to be written practically module by module. The latter provides a degree of extensibility, permitting the model to be scaled to higher or lower levels of organization [24]. In this thesis, we describe experimental and modeling solutions for evaluating one of the most problematic steps in targeted gene and drug delivery, endosomal escape. Through the methods presented here, we provide the characterization of a reporter system for detecting single-endosome escape events, which, together with a rule-based model, give insights into the immense influence of cell type-specific uptake on endosomal escape.

### 1.3 Overview of the thesis

In this thesis, we focus on the interplay between experiment and modeling for understanding uptake and intracellular movement of different NP-sized material, including synthetic gene delivery vectors, mesoporous NPs that can be used as drug carriers, and viruses. In Chapter 2, we introduce the recurring theme of this work, which is understanding the influence of vector uptake and intracellular distribution on individual cell behavior. In the first work, we study synthetic gene delivery; here, we use a mathematical model to show that the high variability of individual cell expression signatures arises from the stochasticity of vector translocation and activation in the nucleus. We also show how the model can predict important vector performance markers, such as percentage of transfection and the number of efficient plasmids per vector complex. We then discuss scalability limitations of the model when processes upstream of nuclear entry are included, such as vector complexation and movement across cellular compartments. In Chapter 3, we present theoretical frameworks, namely pi calculus and the compartment-based PABM calculus, for addressing these problems. We then demonstrate the potential of these frameworks in Chapter 4 by modeling the intracellular distribution of gold nanoparticles (NPs) based on some transmission electron microscopy (TEM) data and parameters derived from detailed studies of gold NP behavior. We also extended the experimental setup to include real-time analysis of individual, light-induced endosomal escape using a nanoparticle (NP)-based detector/delivery system. The experiments and model extensions required to analyze the distribution results that we obtained are detailed in Chapter 5. Finally, in Chapter 6, we present a mechanistic model, formulated in PABM and simulated in the model checker PRISM, that shows how a minimum number of influenza hemagglutinin (HA) can form a fusion pore that facilitates endosomal escape. Although it is currently not possible to

engineer NP carriers that have preassembled pore-forming units, or a fixed numbers of attached molecules, such information may eventually be used to improved non-viral vector performance [25, 26].

## Chapter 2

# Modeling non-viral gene transfer using a hybrid stochastic model

The development of safe and efficient transgene and drug delivery methods has been the focus of extensive research efforts for more than 30 years [27, 28, 29]. Most of the vectors are modified viruses, which have the inherent ability to transfer genetic material to cells; however, safety issues and manufacturing issues have shifted the focus on non-viral vector development [27]. However, the most-used synthetic protocol, lipofection has only been used in 6.4 % of the Phase III vector clinical trials [30, 31]. This is because of the relative inefficiency of the non-viral vector compared to its viral counterparts [27]. At the systemic level, this is due extensive interaction of the vectors with serum proteins, which results in rapid clearance [32]; at the cellular level, this can be due to its inefficient endosomal escape, cytosolic degradation or trapping, and suboptimal transport and nuclear entry [33]. Even if the vector enters the nucleus, the transgene is not necessarily stably integrated into the host genome [29]. Nonetheless, non-viral vectors still present several advantages over viral counterparts, which make further research in the field relevant. For instance, non-viral vectors are associated with less safety, toxicity and immune response-related issues [34]. Additionally, it can carry larger transgenes than 8.0 kb limit of most viral vectors [35]. Currently, a lot of effort is in hybrid vector design. These hybrids ideally contain functional elements from viruses to improve targeting, escape and gene integration efficiency without inducing adverse effects [34]. Nonetheless, the success of any hybrid vector design depends on how well the molecular mechanisms governing its delivery is understood.

It is consequently of interest to come up with a way to systematically evaluate and improve the performance of existing vectors and to predict the behavior of new ones. For this, we need to understand the intracellular processes in which both viral and nonviral gene expression systems are involved, to identify events that control transgene expression, and to identify particular strengths and weaknesses of different vectors that have been tested so far [36]. A number of mathematical and computational models have been constructed in attempts to address this need [36]. In a first approach, we evaluate the performance of two synthetic, DNA-condensing delivery systems, polyethyleneimine (PEI) and lipofectamine, using a simple hybrid mathematical model to describe gene expression characteristics in

individual lung epithelial cells.

## 2.1 Single cell expression data-derived distributions of exogenous gene expression characteristics

Until early 2000, most experiments investigating gene expression were based on population averages. While such data could be used as indicators for transgene expression efficiency, these naturally do not give insight into the expression behavior of single cells. With the development of semi-automated fluorescence microscopy methods, single-cell image tracking and analysis algorithms, and the and a wider range of fluorophores, single-cell time course studies have become feasible, high-throughput sources of data.

A question that we wished to answer using this approach was how steps in artificial gene delivery influence the population expression characteristics and vice versa. In this study, we analyzed exogenous gene expression from two synthetic delivery vectors, PEI, a cationic polymer; and lipofectamine, a lipid-based system delivery system that condense plasmid DNA into nanoparticle-sized complexes. The two vectors were tested on the lung epithelial cell line, BEAS2B, in both unsynchronized and synchronized setups with plasmids encoding enhanced green fluorescence protein (EGFP) as the load. Each experiment was performed for 30 hours, with a 10-minute sampling interval over 25 view fields. From these images, the change in integrated density over time were extracted for single cells; the integrated density measurements calibrated against EGFP-coated beads to yield the approximate number of EGFP molecules per cell. Typical time-course results for PEI and lipofectamine demonstrate response heterogeneity in terms of exogenous gene expression levels and expression onset (Figure 2.1).

To generate distributions of the maximum expression level  $I_{max}$ , the time to half-maximum expression ( $t_{1/2}$ ), and the characteristic rise time ( $t_{rise}$ ), we used the following phenomenological fit function to fit individual time traces:

$$I(t) = \frac{I_{max}}{2} \left[ 1 + \tanh \left( \frac{t - t_{1/2}}{t_{rise}} \right) \right] \quad (2.1)$$

The expression onset time, ( $t_0$ ), was subsequently approximated as  $t_{1/2} - t_{rise}$ . We surmised that the  $I_{max}$  distribution, which reflects steady-state exogenous protein expression, will also be indicative of the number of successfully incorporated plasmids. Distributions of these values are shown in Figure 2.2. Note that lipofectamine has an earlier expression peak at eight hours, compared to the 16 hours required by PEI in non-synchronized cultures; this difference disappears with synchronization, indicating that delivery is cell cycle dependent. Synchronization also results in twofold higher  $I_{max}$  values for both vectors.

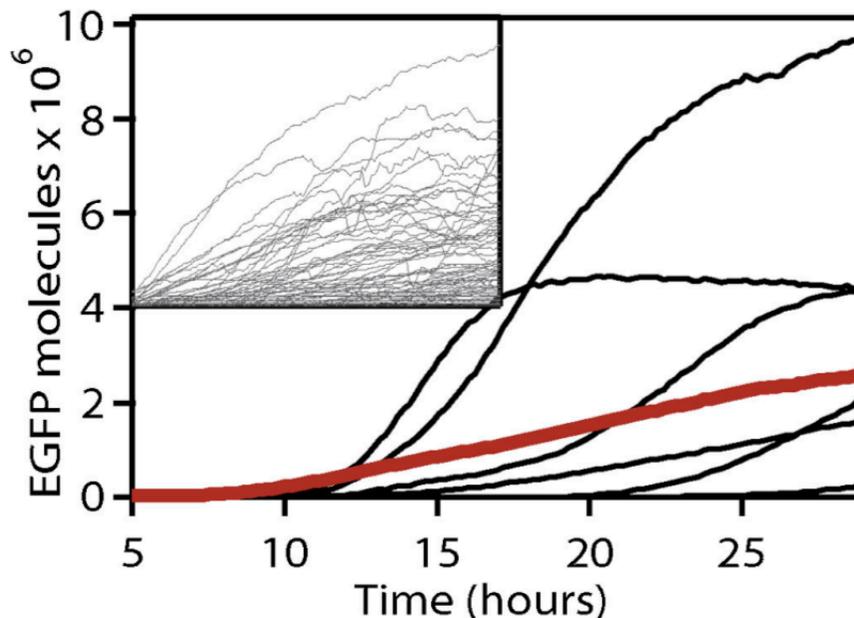


Figure 2.1: Single-cell time traces of exogenous EGFP expression using PEI and lipofectamine (inset). The population average (red) is shown in comparison. Note the loss of information on the expression variability, as well as the sigmoidal shape of the expression curves. Figure from Paper A [18].

## 2.2 Use of a hybrid deterministic/stochastic model to describe delivery processes

To analyze the expression steady states shown in Figure 2.2, we constructed a hybrid stochastic/deterministic model of nuclear delivery and gene expression. First, we attempt to model the EGFP expression steady state; for this, we used a series of differential equations to describe the expression of EGFP from a single, activated plasmid that considers synthesis ( $s_a, s_p$ ) and degradation ( $\delta_R, \delta_U, \delta_G$ ) rates for RNA, unfolded and folded GFP, as well as a maturation rate for folded GFP ( $k_m$ ), from which the following analytical steady state relation can be derived:

$$I_{max} = G(t \rightarrow \infty) = \frac{k_m s_p s_A}{\delta_G \delta_R (k_M + \delta_G)} X \quad (2.2)$$

which can be alternatively expressed as follows:

$$[GFP] = k_{exp} [plasmids] \quad (2.3)$$

Here,  $k_{exp}$  corresponds to the number of proteins expressed in the steady state per plasmid, denoted the expression factor. Interestingly,  $k_{exp}$  is approximately  $4 \times 10^6$  molecules/plasmid,

which leaves  $[plasmids]$  to have a value of 1. This result implies that expression level variance originates from a small number of delivered plasmids. We subsequently model delivery as a two-step stochastic process, where the nuclear translocation step, occurs with a probability  $\mu$  and the plasmid activation step occurs with probability  $q$ . The activation step refers to the actual unpacking of the plasmid from a vector and its subsequent expression. We assumed that nuclear delivery occurs rarely, and that the events are statistically independent:

$$P(C) = \frac{\mu^C}{C!} e^{-\mu} \quad (2.4)$$

Supposing that there are  $X$  activated plasmids, then  $n \geq X$  plasmids must have first entered the nucleus, with each of the  $n$  plasmids having a probability  $q$  of being activated. The probability  $P(X)$  of finding  $X$  activated plasmids is thus given by a binomial distribution:

$$P(X|n) = \binom{n}{X} q^X (1-q)^{n-X} \quad (2.5)$$

The distribution of expressed plasmids, directly related to the number of expressed GFPs, can then be related to equation 2.3 with a well defined mean:

$$\langle [plasmids] \rangle = \mu m q \quad (2.6)$$

The transfection ratio (TR) is then related to this mean, and depends on the average number of complexes delivered, as well as  $\bar{q}$ , the effective probability that from *any* complex, *at least* one plasmid is transcribed:

$$TR(\mu, m, q) = 1 - \exp\{-u\bar{q}\} \quad (2.7)$$

For our data, TR is approximately 20 % (Figure 2.3), while the average number of activated plasmids is three, confirming previous reports and estimates on the extremely low amount of vectors that enter the nucleus [19]. The model fits data from synchronized cells; for non-synchronized cultures, the fits are less ideal, given the occurrence of noise occurring from cell cycle-dependence of gene expression.

### 2.3 Limitations of differential equation models

The results of the study highlight the effect of vector loss, and low nuclear entry and activation probabilities. Of the processes upstream nuclear entry, poor endosomal escape has been identified as the most likely bottleneck in the delivery process [27]. It is thus of interest to develop a method for observing single endosomal escape events, which could be used to directly evaluate vector loss. It would likewise be necessary to extend the model to handle the topological change arising from endosome lysis and details pertaining to its necessary triggers. In the next chapters, we proceed by describing rule-based modeling

solutions (Chapter 3) and an application (Chapter 4) that demonstrates its adequacy and suitability for modeling individual endosome lysis events (Chapter 5).

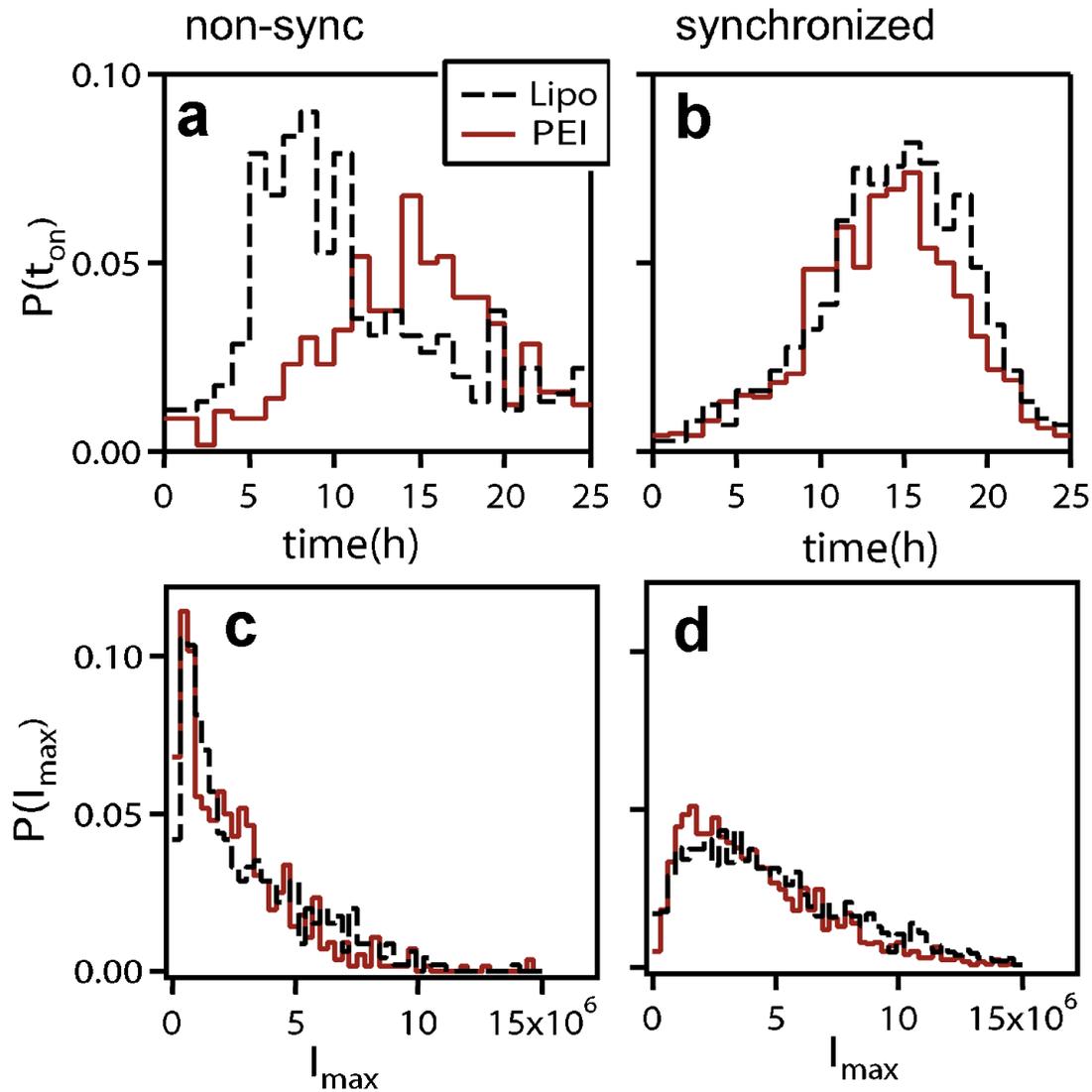


Figure 2.2: EGFP expression statistics for PEI- (red) and lipofectamine- (black, dashed) mediated transfection in synchronized and non-synchronized cultures for the expression onset time,  $t_{on}$  (a and b) and maximal EGFP values,  $I_{max}$ , showing strong expression variability, even after synchronization. Note the cell cycle dependence of of transfection. Figure from Paper A [18].

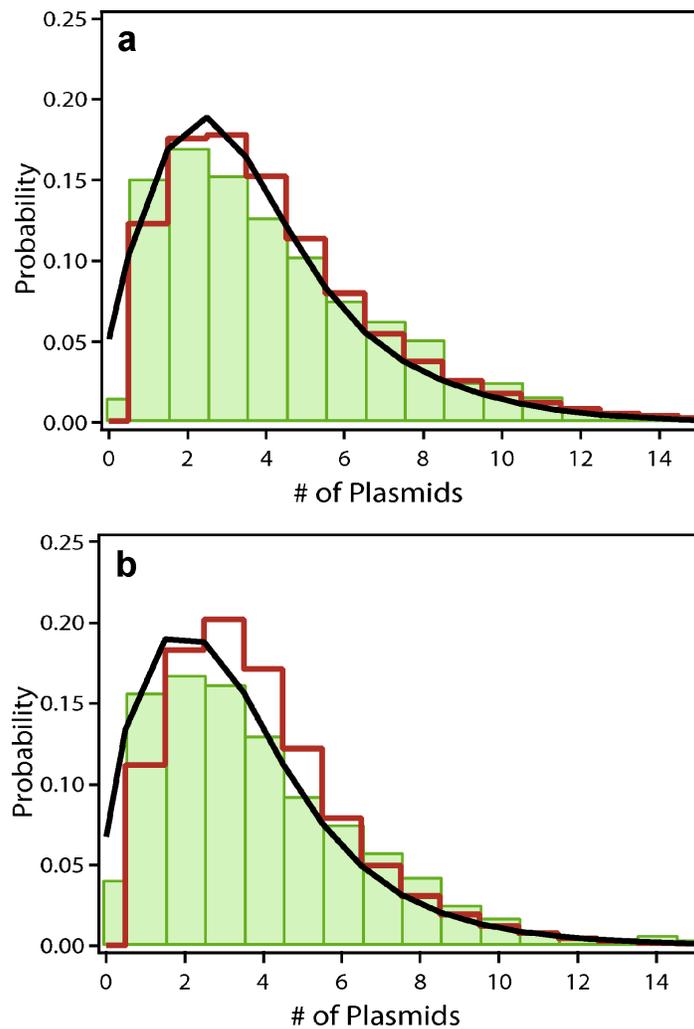


Figure 2.3: A superposition of the model result (black) with the distribution of expressed plasmids (red) and the protein expression, scaled down by the average number of GFP molecules per plasmid at steady state (green) for PEI (a) and lipofectamine (b) show a 20% transfection efficiency for both vectors. The graph also shows that the average number of activated plasmids is three. Figure from Paper A [18].



## Chapter 3

# Modeling stochastic cellular processes: dealing with compartments and complex networks

The gene delivery model that we have presented is able to predict the percentage of transfected cells, the cotransfection ratio, the distribution of expression levels, and the number of efficient plasmids per complex, which may be used for evaluating gene delivery vectors [18]. However, if the scopes of both experiment and model are extended to include mechanistic details of endosomal escape or cytoplasmic trafficking, then amount of detail that has to be incorporated makes it more difficult to create the model using ODEs. For instance, an approach for improving the cytoplasmic transport of vectors could require a study of viral components that mediate engagement and trafficking along microtubules. One might also need to look at signal transduction pathways that viruses activate to induce trafficking [37, 38]. Putting all these information in a model could yield insights into the combinations of components that one could put in a vector. Similarly, comparing models of non-viral and hybrid vector behavior would provide information on the transport components and mechanisms that can be exploited better [39], while simultaneously resolving some conflicting information on artificial vector transport [40].

Apart from the complexity that inevitably arises from the inclusion of such details, the fact that all these events occur in different cellular compartments makes the specification or extension of existing ODE models more challenging. Such a problem is likely to be encountered easily when constructing NP cytotoxicity models, where intracellular particle distributions have to be modeled explicitly, alongside important details of cell signaling pathways [41]. In order to address these needs, it is of interest to look at alternative modeling methods, specifically rule-based modeling, and the formalisms within the emerging area of executable biology [42]. In particular, we focus on pi calculus [43]; and the projective brane calculus with activate, bud and mate as primitive actions (PABM) [44], a generalization and extension of brane calculus [45], which handles compartments and membrane operations explicitly.

### 3.1 Introduction to pi calculus

Pi calculus was developed by Milner and colleagues for modeling mobile, concurrent communicating systems [43], allowing the high-level description and logical analysis of interaction, communication and synchronizations between agents known as *processes*. The basic objects, operations and reduction rules of pi calculus are summarized in Table 3.1.

| Object               | Syntax                      | Description  |
|----------------------|-----------------------------|--|
| Process              | $P, Q$                      | Agents/objects that make up a system   |
| Channel              | $?x, !x$                    | Named communication ports of processes; channels are associated with rates, and may be used to pass attributes, channels, or processes between processes |
| Attributes           | $P(\text{att}:\text{type})$ | Process characteristics  |
| Operation            | Syntax                      | Description  |
| Concurrency          | $P \mid Q$                  | $P$ and $Q$ are processes executed concurrently  |
| Summation            | $!x.P + ?y.Q$               | Represents a competitive choice between mutually exclusive interactions  |
| Restriction (Series) | $(vx).P$                    | Creation of a new name in $P$  |
| Replication          | $!P$                        | Associated with processes that can always create a new copy of $P$   |

Table 3.1: Objects and operations of pi calculus

In pi, systems are modeled using multiple processes that communicate with each other in a pairwise, synchronized manner over named channels [43]. A computation step is denoted as  $P \xrightarrow{\text{rate}} P'$ ; the main reduction rule of pi calculus is denoted  $?x(z).P \mid !x(y).Q \xrightarrow{\text{rate}} P \mid Q[z/y]$ , showing name-passing between channels. Alternately, an internal state change may be denoted as  $P \xrightarrow{\tau} P'$ , where  $\tau$  is a *delay* reaction. These features of pi calculus permit process modification through the acquisition of new communication capabilities [46]. A stochastic extension, where non-determinism is replaced by exponentially-distributed race conditions, was added by Priami for performance modeling of distributed systems; the stochastic extension allows pi calculus systems to be quantitatively analysed with its mapping to a continuous time Markov chain [47]. This particular feature has been used extensively in quantitative models of biological systems [45].

The use of stochastic pi calculus as a foundation for biological modeling and stochastic simulations was proposed by Regev, because of the adequacy of its coverage, semantics, and dynamics to capture the molecular organization and dynamic behavior of these systems [46, 24]. In particular, Regev mentions three features of pi which make it a good candidate for biomolecular process abstraction: 1) *computability*, which allows the abstraction to simulate system behavior while providing a qualitative and quantitative reasoning behind the system properties; 2) *understandability*, which presents an ease of correspondence

between biological concepts and objects (Table 3.2); and 3) *extensibility*, which permits the simulation to be scaled to higher levels of organization than molecules [24]. Pi calculus has since been successfully applied in EGFR network modeling [48], early T cell signaling [49], and more recently, in NP-cell interactions (Chapter 4).

| Biochemical object                               | Pi calculus object |
|--|--------------------|
| Reactant (Proteins, ligands, compartments, etc.) | Process            |
| Binding site                                     | Channel            |
| State  | Attribute          |

Table 3.2: Correspondence between biochemical and pi objects. Note that the definition of the process is broad enough to allow it to be used as an abstraction for a compartment.

## 3.2 The Stochastic Pi Machine (SPiM): a pi calculus implementation

### 3.2.1 Introduction to SPiM

The stochastic pi Machine (SPiM) is an implementation of a stochastic pi calculus variant [50]. In SPiM v.0.05, each process is encoded as a computer object with a scope that corresponds to the visibility of its associated channels; that is, a process  $P$  having a hidden channel  $x$  excludes another process  $Q$  from seeing  $x$ . A process may evolve as a result of an action on a channel or through a delay, which is often used to represent a change in internal state of a process [48]. A state change is subsequently effected by changing a parameter specification, the attribution of a new channel, or the redefinition into another process.

In addition to the basic pi calculus implementation, SPiM allows the use of simple programming language-type statements, including basic conditional statements and boolean operations. This effectively extends the rule-based formalism that pi calculus provides (Section 3.2.2). Parameters in SPiM are also not limited to channel and channel passing, but include attributes that may assume different data types. Although not an integral version of CCS, these attributes have been used in order to provide enough flexibility to handle problems ranging from very straightforward handling of state changes to its *ad hoc* use as coordinates in 3D space [51]. A version of SPiM with a graphical user interface has recently been developed; however, this release does not support run automation, for which reason all of the SPiM models that were created for the research were specifically for v.0.05. SPiM uses the Gillespie algorithm (Section 3.2.3) to choose the channel or delay on which a reaction occurs, as well as the duration of this reaction.

### 3.2.2 Rule-based modeling in SPiM

A *rule* specifies features of a reactant that are either required for or are affected by reactions, and how these features are to be “rewritten”, in case the reaction is chosen for execution [52, 53]. Rules can thus be viewed as generalized reaction patterns that permit reactant recognition, a mapping of reactants to products, and an association of a rate law with this mapping [53]. The view of a rule is very local – that is, it often represents a local pattern of a site or a state (Figure 3.1). For instance, given a pair of proteins  $P$  and  $Q$  that may interact only when  $P$  is phosphorylated at site  $x$ , a rule would permit any such reaction, regardless of whether  $P$  or  $Q$  are bound to other proteins or not – thus, the rule only looks at  $x$ , and not the whole  $P$  and  $Q$  – and has the potential to encompass a lot of reactions that would have otherwise been treated as occurring between distinct molecular species [52].

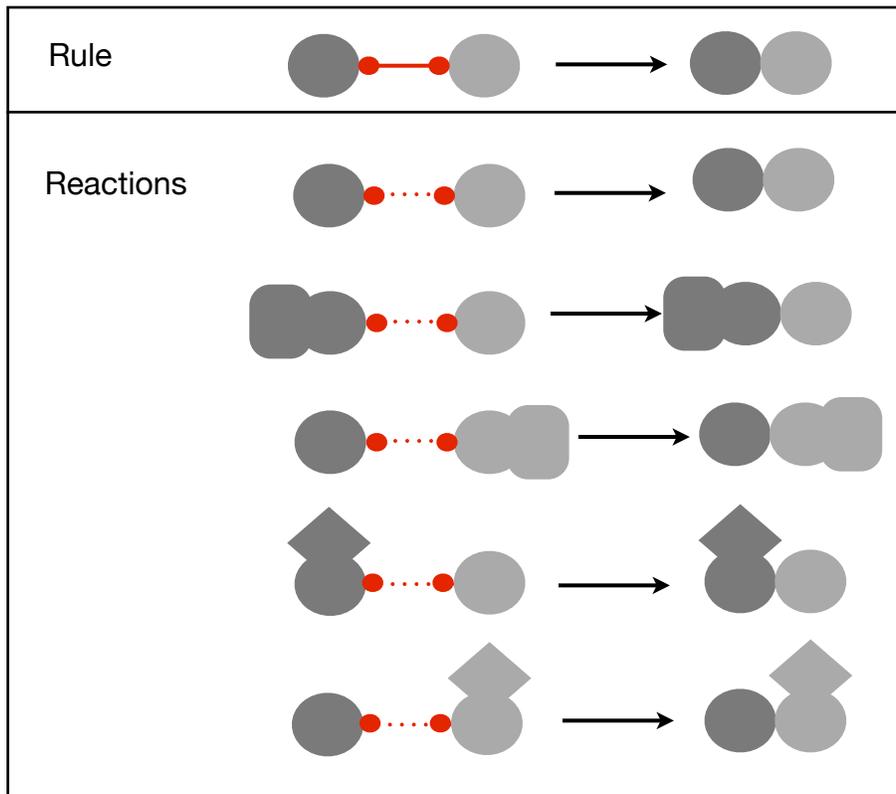


Figure 3.1: Rule based modeling uses patterns to identify possible reactions, and to map all the reactions encompassed by these patterns into products. In this figure, the reactions that can be encompassed by the illustrated rule are shown. By reducing the number of reactions that have to be defined, rule based models can significantly reduce the combinatorial complexity that features frequently in models of biological systems.

The benefits of a rule-based approach is most evident when dealing with multistate

protein modifications, a frequent feature of signaling pathways. Consider the case of epidermal growth factor (EGF) signaling, which involves a network of at least 211 reactions and 322 interactants. Each interactant generally has more than one state that determines its possible binding partners [54]. However, the reactions can be categorized into more general classes such as state transitions, including phosphorylation; association and dissociation; or translocation-type reactions [54]. Assuming that  $n$  of the state transition reactions are phosphorylation reactions, and a general pattern of *phosphorylase* | *substrate*  $\rightarrow$  *substrate* -  $P$ , where *substrate* -  $P$  is a state of the *substrate* process, then these  $n$  reactions can be recognized using a single rule, and would not have to be defined individually.

In addition to rules, SPiM also implements programming language-like constructs such as *if* - *then* statements, which permit the inclusion of reaction preconditions on top of patterns. These constructs could also be used for defining rules in an alternative manner. This is particularly advantageous for problems that could not be represented as chemical reaction-type models.

The strength of combining some logic-based operations of programming languages with a rule-based modeling becomes evident when we need to define thresholds. Suppose that  $P$  has two attributes,  $a$  and  $b$ , and  $Q$  interacts with  $P$  with a rate  $f(a)$  if  $a \leq b$  and  $f(b)$  if  $b < a$ . A SPiM code representation is very straightforward:

```
let P(a:int, b:int) =
  if a <=b then
  delay@f(a); !react
  else
  delay@f(b); !react
    and Q() = ?react; PQ()
    and PQ() = ()
```

Assuming further that  $a$  and  $b$  can each range over values  $1..n$  and  $1..m$ , respectively, a plain Gillespie implementation that is able to track all reactions of  $P(a : 1..n, b : 1..m)$  will result in a reaction matrix where the  $n * m$  reactions of  $P$  have to be explicitly defined. Overall, these features of SPiM eliminate redundancies often required when ODEs, or even a direct implementation of Gillespie, are used.

### 3.2.3 The Gillespie algorithm

The stochastic simulation algorithm (SSA) of Gillespie addresses the general question of the time evolution of  $N$  reactive chemical species in a fixed volume  $V$ , given the numbers of each of the  $N$  species, as well as their  $M$  reaction channels. More specifically, it was conceived to describe *fluctuations* in the molecular population levels of reacting systems that cannot be captured by reaction rate equations [55]. In SSA, reaction constants are treated as probabilities than rates, and the temporal behavior of the system is a Markovian random walk in  $N$  dimensions of the  $N$  species. The time evolution of the probability  $p(x, t)$  that the system will be in state  $x$  at time  $t$  is described by the *master equation*.

The physical basis for the stochastic formulation of chemical kinetics lies in the premise that collisions in a well-stirred system in equilibrium occur essentially at random. For such a system, it is possible to calculate the *probability*  $P(R_{1,2})$  of all pairwise collisions 1, 2, occurring in  $V$  within an infinitesimal time interval  $[t, t + \delta t]$ :

$$P(R_{1,2}) = X_1 X_2 V^{-1} \pi r_{12}^2 \bar{v}_{12} \delta t \quad (3.1)$$

where  $X_1$  and  $X_2$  are the concentrations of molecules  $S_1$  and  $S_2$ , respectively;  $r_{12}$  is the sum of the radii of  $S_1$  and  $S_2$ ; and  $\bar{v}_{12}$  is the average relative velocity of the molecules, assumed to follow a Maxwellian velocity [55]. It is further asserted that a *stochastic reaction constant*  $c_1$ , solely dependent on the physical properties of the reacting molecules and the temperature of the system, exists, such that the probability of the collision within the time interval  $[t, t + \delta t]$  is:

$$P(R_{1,2}) = X_1 X_2 c_1 \delta t \quad (3.2)$$

This can be generalized for chemical species  $S_i, i = 1..N$  that react through  $M$  reaction channels  $R_\mu, \mu = 1..M$ , for which there must be  $M$  constants  $c_\mu, \mu = 1..M$ . These yield an expression for the *average* probability that a particular combination of  $R_\mu$  reactants will react within  $[t, t + \delta t]$ :

$$\bar{P}(R_\mu) = c_\mu \delta t \quad (3.3)$$

The next step would be actually describe the stochastic time evolution of the system, which is traditionally done by setting up the master equation – in this case, directly derivable from the sum of probabilities of the  $M + 1$  ways by which the system can arrive at a state  $(X_1...X_N)$  at time  $t + \delta t$ :

$$P(X_1, \dots, X_N; t + \delta t) = P(X_1, \dots, X_N; t) \left[ 1 - \sum_{\mu=1}^M a_\mu \delta t \right] + \sum_{\mu=1}^M B_\mu \delta t \quad (3.4)$$

where  $a_\mu \delta t$  is the probability that a reaction  $R_\mu$  will occur, given the state of the system at time  $t$ :

$$a_\mu \delta t \equiv c_{\mu u} \delta t \times R_\mu C \quad (3.5)$$

Here,  $R_\mu C$  is the number of distinct  $R_\mu$  molecular combinations in the state.  $B_\mu \delta t$  in Equation 3.4, on the other hand, gives the probability that the system is one reaction  $R_\mu$  removed from the state  $(X_1...X_N)$  at time  $t$ , and then undergoes the reaction  $R_\mu$  at  $t + \delta t$ .

However, *solutions* to such equations are intractable. In its stead, Gillespie proposed the *simulation* of this time evolution, which answers the two critical questions to make the system move forward in time: (1) when will the next reaction occur? and (2) what kind (or which) reaction will it be? To this end, a function  $P(\tau, \mu) \delta t$ , which gives the probability of the *next*  $R_\mu$  reaction occurring within the time interval  $[t + \tau, t + \tau + \delta t]$  was defined;  $\tau$  and  $\mu$  answer the two questions posed previously. To assign  $\tau$  and  $\mu$ , the following steps are taken:

1. For each  $R_\mu$ , define the hazard function,  $h_\mu$ :

(a) For reactions of the form  $S_1 + S_2 \rightarrow Product$ :

$$h_\mu = X_1 X_2 \quad (3.6)$$

(b) For reactions of the form  $2S_1 \rightarrow Product$ :

$$h_\mu = \frac{1}{2} X_1 (X_1 - 1) \quad (3.7)$$

2. Calculate  $P(\tau, \mu)$  from the product of  $(P_0(\tau))$ , the probability that no reactions will occur within  $[t + \tau, t + \tau + \delta t]$  state of the system at time  $t$ ; and  $a_\mu \delta t$ , which is the probability that reaction  $R_\mu$  occurs:

$$P(\tau, \mu) \delta t = P_0(\tau) a_\mu \delta t \quad (3.8)$$

Noting that the probability that no reaction will occur at  $\delta t$  is  $1 - \sum \nu a_\nu \delta t'$  within time  $\delta t'$ , then:

$$P_0(\tau + \delta t') = P_0(\tau') \left[ 1 - \sum_{\nu=1}^M a_\nu \delta t' \right] \quad (3.9)$$

Integration of which yields:

$$P_0(\tau + \delta t') = \exp \left[ - \sum_{\nu=1}^M a_\nu \tau \right] \quad (3.10)$$

so that Equation 3.8 can be rewritten as:

$$P(\tau, \mu) \delta t = \begin{cases} a_\mu \exp - a_0 \tau, & 0 \leq \tau < \infty, \mu = 1 \dots M \\ 0 & \end{cases} \quad (3.11)$$

where  $a_\mu \equiv h_\mu c_\mu$  and  $a_0 \equiv \sum_{\mu=1}^M a_\nu \equiv \sum_{\mu=1}^M h_\nu c_\nu$ . It is noteworthy to mention that  $P(\tau, \mu)$  depends on all reaction constants, as well as the count of all reactant species [55]. With this background, the stochastic simulation algorithm can be algorithmically defined as follows:

1. Initialize  $t = t_0$ , the system state  $X = X_1 \dots X_N$ , and specify the reaction constants  $c_\mu$ ,  $\mu = 1 \dots M$ . Set the reaction counter to 0.
2. Calculate the values of  $a_1 = h_1 c_1 \dots a_M = h_M c_M$  (Equations 3.6 and 3.7), as well as the value of  $a_0$ , given by the sum of  $a_\nu$  values

3. Generate two random numbers,  $r_1$  and  $r_2$ , taken from a unit interval
4. Generate values for  $\tau$  and  $\mu$  as follows:

$$\tau = \frac{1}{a_0 x} \ln \frac{1}{r_1} \quad (3.12)$$

$$\mu = \sum_{\nu=1}^{\mu-1} a_\nu < r_2 a_0 \leq \sum_{\nu=1}^{\mu} a_\nu \quad (3.13)$$

5. Update the time as  $t + \tau$ ; update the system state, and recalculate the reaction hazards for species whose states have changed. Increase the counter by 1.
6. Return to step 2, and iterate until  $t \geq t_{end}$

Several methods exist for the Gillespie algorithm, including the direct method [55], and subsequent alternative solutions, which attempt to improve the performance of the algorithm such as the next reaction method [56], the slow-scale stochastic simulation algorithm [57], and tau-leaping [58, 59]. SPiM v.0.05 implements the direct method [50].

### 3.3 Brane calculi: modeling stochastic membrane processes

While pi calculus provides a possibility of representing compartments as processes (Table 3.2) between which other processes or channels are passed, the fact that it does not have specific reduction rules for these compartment processes could limit its expressivity, given that the consequences of a compartment transformation is not limited to itself, but extends to its contents – in this case represented as attributes. As an example, we can take a fusion-type reaction between two compartments modeled as processes that have different contents. This can be roughly modeled in SPiM using the following commands:

```
let Compartment_A (a:Proc) =
!fuse(a)
and Compartment_B (b:Proc) =
?fuse(a'); (Compartment_AB (a, b) )
and Compartment_AB(a:Proc, b:Proc) = ()
```

Now, supposing `Compartment_AB` itself can merge with other compartments, the code has to be modified as follows:

```
let Compartment_A (a:Proc) =
!fuse(a)
and Compartment_B (b:Proc) =
```

```

?fuse(a'); (Compartment_AB (a, b) )
and Compartment_AB(a:Proc, b:Proc) =
!fuse(a,b)
and Compartment_X(x1:Proc ... xn:Proc) = ()
?fuse(a',b'); Compartment_ABX(a,b,x1 ... xn)
and Compartment_ABX(a: Proc ,b: Proc ,x1 ... xn: Proc) = ()

```

One problem that becomes evident in this example is that it is not possible to create a single rule for fusion. This limitation arises from the fact that pi, and particularly SPiM, semantics neither permit the use of process list structures as attributes, nor have native operators that tell the machine what to do with this list. Consequently, a pi implementation that involves topological operations on compartment processes with *deeply nested* content suffers from a degree of combinatorial complexity.

Brane calculus was introduced for explicitly handling these effects of membrane-bound compartments, and more specifically, of the topological changes that these can undergo, in models of biological systems [45]. Cardelli's brane calculus introduces the concept of *branes* as dynamic wrappers that contain other branes or (pi) processes. *Actions* refer to the rules that define how branes may interact, and the consequences of these reactions; these are directly inspired by the actual topological transformations that biological membranes undergo. Actions include uptake-type reactions in the form of *phago* and *pino*; fusion-type reactions in the form of *mate* and *exo*; and split-type reactions in the form of *bud* and *drip*. Figure 3.2 summarizes the rules of the calculus, while Tables 3.3 and 3.4 provide details of the syntax and reduction rules. Note that the fusion-type reactions, *exo* and *mate* result in the mixing the branes, or of both the branes and their content, respectively.

| Object            | Syntax  | Description  |
|-------------------|---|--|
| System            | $[\sigma][P]$   | Parallel compositions of Branes and their contents   |
| Brane             | $[\sigma]$  | Membrane bound compartment that can contain other Branes or (pi processes)                                   |
| Action            | $\sigma, \tau \in \{\text{phago, pino, exo, mate, bud, drip}\}$ | Names of brane calculus rules; each action is associated with a channel to delineate interaction specificity |
| Process           | $P, Q$  | Agents/objects that  |
| Operations        | Syntax  | Description  |
| Parallel, Systems | $[\sigma][P] \circ [\tau][Q]$                                   |  |
| Parallel, Actions | $\sigma \tau$   |  |
| Series            | $\tau.\sigma$   |  |
| Choice            | $\sigma + \tau$   |  |

Table 3.3: Objects and operations of Brane calculus

As indicated previously, pi processes could be incorporated in a brane model as brane contents, with its availability being restricted by the brane topology, thus providing a

| Action | Reduction rule  |
|--------|---|
| phago  | $[\sigma \text{phago!}x \langle P \rangle] \circ [\tau \text{phago?}x \langle Q \rangle] \rightarrow [\tau \langle [ - ] \langle [\sigma \langle P \rangle] \rangle Q \rangle]$ |
| pino   | $[\sigma \text{phago!}x \text{phago?}x \langle Q \rangle] \rightarrow [\sigma \langle [ - ] \langle [ - ] \rangle Q \rangle]$   |
| exo    | $[\tau \text{exo?}x \langle [\sigma \text{exo!}x \langle P \rangle] Q \rangle] \rightarrow P \circ [\tau \sigma \langle Q \rangle]$   |
| mate   | $[\sigma \text{mate!}x \langle P \rangle] \circ [\tau \text{mate?}x \langle Q \rangle] \rightarrow [\sigma \tau \langle P \circ Q \rangle]$                                     |
| bud    | $[\tau \text{bud?}x \langle [\sigma \text{bud!}x \langle P \rangle] Q \rangle] \rightarrow [\tau \langle [\sigma \langle P \rangle] Q \rangle] \circ [\tau \langle Q \rangle]$  |
| drip   | $[\sigma \text{bud!}x \text{bud?}x \langle Q \rangle] \rightarrow [\sigma \langle [ - ] \rangle] \circ [\sigma \langle Q \rangle]$  |

Table 3.4: Reduction rules of Brane calculus

language that sufficiently captures most features of a cellular system. Going back to the example given at the start of this section, the whole can be re-written simply in brane as follows; note that there are no explicit specifications of the product, given that this would be resolved by the reduction rules, with the order of reactions being determined as in Section 3.2.3:

$$[\text{mate!}x|\text{bud?}x|\text{bud!}x \langle A [ - ] \langle a \rangle \rangle] \circ [\text{mate!}x \langle B [ - ] \langle b \rangle \rangle] \quad (3.14)$$

Danos and Pradali er extended the calculus by introducing the concept of *sidedness* – that is, that the outer and inner layers of a biological membrane can be specified to have different interaction capabilities – in order to create a language closer to biological membranes (Figure 3.3, [60]). The calculus essentially has the same objects and actions as Cardelli’s brane calculus, and has the difference in the syntax that reflects the sidedness of a brane:  $[\langle \sigma; \tau \rangle \langle P \rangle]$ , with  $\sigma$  representing the actions on the outer leaflet, and  $\tau$ , of the inner leaflet.

### 3.4 A Projective Brane calculus with activate, bud and mate as primitive actions (PABM)

In our work [44], we further extended projective brane calculus by introducing the concept of *domains*, or brane patches. Unlike in previous calculi where branes are treated as continuous objects, domains allow portions of membrane to behave autonomously (Tables 3.5 and 3.6). The use of domains imbues greater control over how membranes are merged and split. This is most evident in the case of bud, where the extruded membrane is derived exclusively from the domain that received the trigger to bud (Figure 3.4).

Additionally, it could be use for more detailed modeling of the molecular events that govern the topological changes in the first place. In recent research, it was shown that transport vesicle formation may be driven by the steric pressure arising from membrane protein interactions; more specifically, asymmetrically distributed *luminal* proteins within the vesicle could alter the physical properties of a membrane and influence its curvature [61]. In this particular example, the use of domains to represent the asymmetric distribution of the proteins, and to use sidedness to represent the fact that these proteins are in the

| Object                | Syntax  | Description   |
|-----------------------|---|---|
| System                | $[\vec{\rho}](P) \circ Q$   | System of Branes  |
| Brane                 | $[\vec{\rho}]$  | Membrane bound compartment that can contain other Branes or (pi processes)                                      |
| Domain                | $\vec{\rho} \equiv \langle \sigma_1; \sigma_2 \rangle$                | Discrete part of a membrane containing functionally-related actions   |
| Action                | $\sigma \in bx, mx, !x$   | ; reduced to mate and bud, with its corresponding activator   |
| Process               | $P, Q$  | Agents/objects that make up a system  |
| Operations            | Syntax  | Description   |
| Parallel, Systems     | $[\vec{\rho}](P) \circ [\vec{\rho}](Q)$                               | Systems $[\vec{\rho}](P) \circ [\vec{\rho}](Q)$ occur concurrently  |
| Parallel, Domains     | $\langle \sigma_1; \sigma_2 \rangle   \langle \tau_1; \tau_2 \rangle$ | Domains $\langle \sigma_1; \sigma_2 \rangle   \langle \tau_1; \tau_2 \rangle$ occur concurrently                |
| Parallel, Actions     | $\sigma, \tau$  | Actions $\sigma$ and $\tau$ occur concurrently  |
| Parallel, Processes   | $P \mid Q$  | $P$ and $Q$ are processes executed concurrently   |
| Series                | $\sigma.\tau$   | Indicates that $\tau$ can only be executed after $\sigma$   |
| Choice                | $\sigma + \tau$   | Indicates that either $\sigma$ or $\tau$ can occur; selection of one results in the elimination of other option |
| Replication, Parallel | $(n)\sigma, (n)\vec{\rho}$  | $n$ instances of action $\sigma$ or domain $\vec{\rho}$   |
| Replication, Series   | $\sigma^{(n)}$  | $n$ instances of $\sigma$ executed one after the other  |

Table 3.5: Objects and operations of PABM calculus

interior would be very natural:

$$[bx.(n)\langle -; Protein_A \rangle | !x.(m)\langle -; Protein_B \rangle ](ER) \quad (3.15)$$

Here,  $n$  and  $m$  represent the critical numbers of  $Protein_A$  and  $Protein_B$ , accumulated on separate domains, that have to be present before a transport vesicle from the endoplasmic reticulum (ER) is formed with bud.

Apart from extending brane, we also provide a generalization of its rules, which we reduce from six to two. This is possible by using the concepts of projective invariance and projective equivalence, which first appeared in projective brane calculus [60]. Briefly, projective invariance postulates that the *physics* of an interaction does not make a distinction between inside and outside. Given the following system:

$$[\langle -; mx \rangle ]( [ !x; - \rangle ](P) \circ [\langle mx; - \rangle ](Q) \circ S) \quad (3.16)$$

Assuming that  $S$  is projected to infinity, Eq. 3.16 becomes transformed as follows,

| Action              | Reduction rule  |
|---------------------|---|
| bud (Type I), $b$   | $P \circ [\vec{\rho}_1   \langle \sigma_2; bx \rangle ] \llbracket [\vec{\rho}_2   \langle !x; \sigma_2 \rangle ] \llbracket Q' \rrbracket \circ R \rrbracket \longrightarrow$<br>$P \circ [\langle \sigma_1; - \rangle ] \llbracket [\vec{\rho}_2   \langle -; \sigma_2 \rangle ] \llbracket Q' \rrbracket \rrbracket \circ [\vec{\rho}_1 ] \llbracket R \rrbracket$ |
| bud (Type II), $b$  | $[\vec{\rho}_1   \langle bx; \sigma_1 \rangle ] \llbracket P \rrbracket \circ [\vec{\rho}_2   \langle !x; \sigma_2 \rangle ] \llbracket Q' \rrbracket \circ R \longrightarrow$<br>$[\vec{\rho}_1 ] \llbracket P \circ [\langle \sigma_1; - \rangle ] \llbracket [\vec{\rho}_2   \langle -; \sigma_2 \rangle ] \llbracket Q' \rrbracket \rrbracket \rrbracket \circ R$ |
| mate (Type I), $m$  | $P \circ [\vec{\rho}_1   \langle !x; \sigma_1 \rangle ] \llbracket Q \rrbracket \circ [\vec{\rho}_2   \langle mx; \sigma_2 \rangle ] \llbracket R \rrbracket \longrightarrow$<br>$P \circ [\vec{\rho}_1   \vec{\rho}_2   \langle -; \sigma_1, \sigma_2 \rangle ] \llbracket Q \circ R \rrbracket$   |
| mate (Type II), $m$ | $[\vec{\rho}_1   \langle \sigma_1; mx \rangle ] \llbracket P \circ [\vec{\rho}_2   \langle !x; \sigma_2 \rangle ] \llbracket Q \rrbracket \rrbracket \circ R \longrightarrow$<br>$[\vec{\rho}_1   \vec{\rho}_2   \langle \sigma_1; \sigma_2 \rangle ] \llbracket P \rrbracket \circ Q \circ R$  |

Table 3.6: Reduction rules of PABM

effectively transmuted inside to outside:

$$[\langle mx; - \rangle ] \llbracket [\langle !x; - \rangle ] \llbracket P \rrbracket \circ [\langle mx; - \rangle ] \llbracket Q \rrbracket \circ S \rrbracket \quad (3.17)$$

Projective equivalence is defined as the least equivalence relation that produces this transmutation; using this principle, mate and bud can be viewed as inverse actions, resulting in the generalization of Brane reduction rules from six to mate- and bud-type reactions in PABM (Figure 3.5, Table 3.5). Note that the notation, and the subsequent pairing of actions, was also changed. While previous calculi required the strict pairing *both* actions and channels, such as  $\text{bud}!x/\text{bud}?x$  and  $\text{mate}!x/\text{mate}?x$ , PABM uses just the channel name and type for a pairing, so that  $bx$  and  $mx$  can both be executed by a single  $!x$  trigger. As with pi calculus, PABM will use the Gillespie algorithm for choosing the next reaction. Currently, an executor for PABM is being developed by an external collaborator using the term-rewriting language MAUDE. Further details of features and possible applications of PABM are discussed in Paper B.

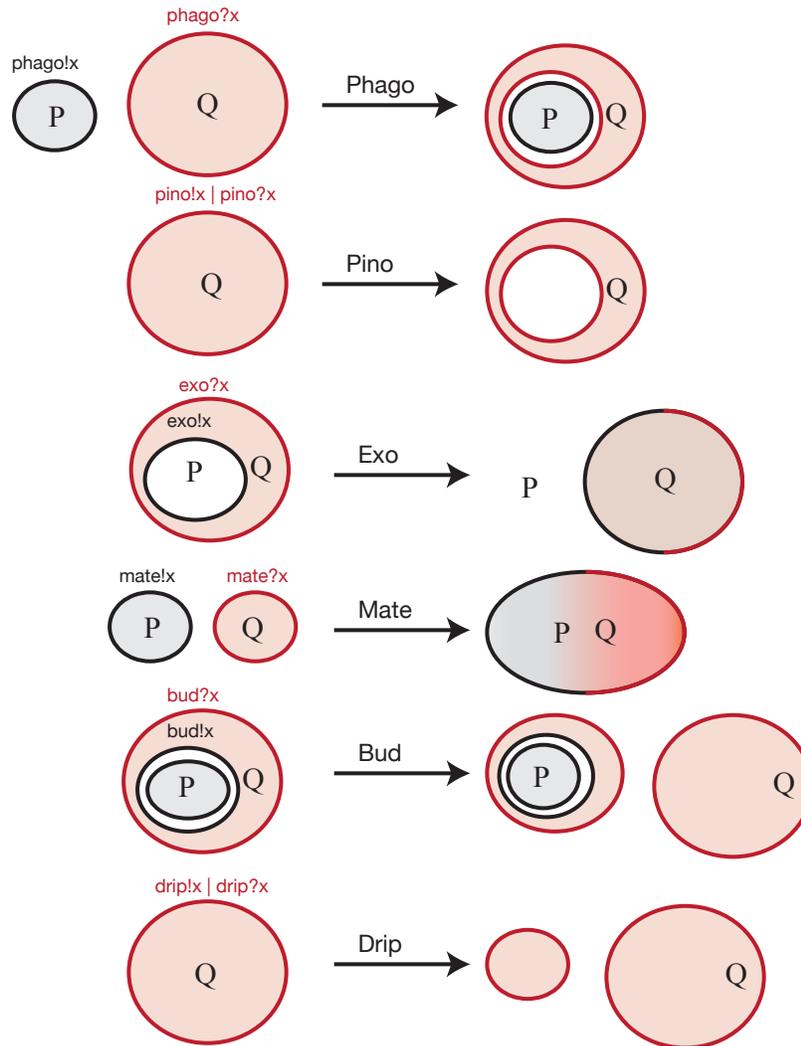


Figure 3.2: Graphical representation of the reduction rules of Cardelli's brane calculus, which generalize the topological changes of biological membranes. Sender (black) and receiver (red) branes and their respective contents are colored to demonstrate the origin of newly-created branes in the product, as the case is in *phago*, *pino*, *bud* and *drip*; or brane or content mixing, as seen in *exo* and *mate*. Figure adapted from [45].

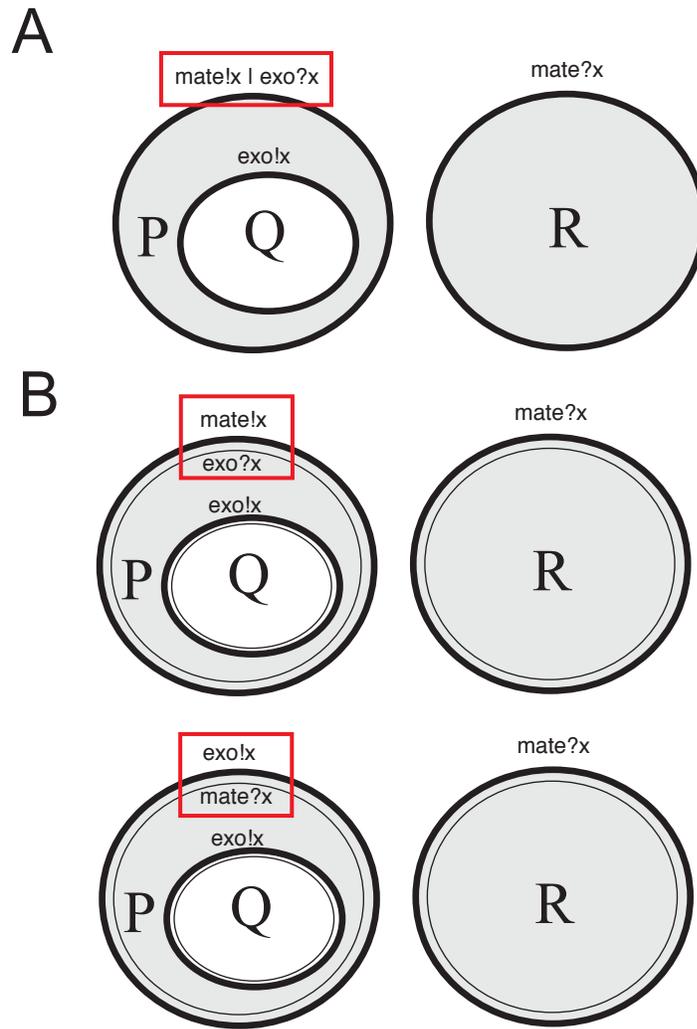


Figure 3.3: Differences between Brane (A) and Projective Brane (B) Calculi. The use of Projective Brane Calculus provides richer semantics that incorporate the effect of protein insertion onto membranes. In this example, a brane calculus system (A) could be more unambiguously expressed (B); note that the two options in (B) will have very different system evolution paths, with the second option being a terminal state.

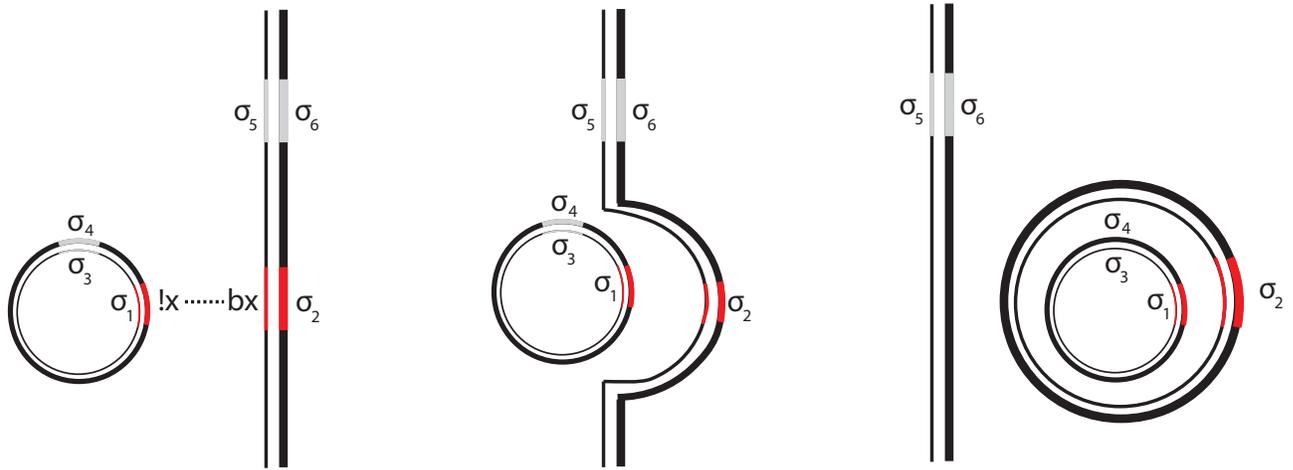


Figure 3.4: PABM domains delineate discrete regions of a brane on which operations can be performed autonomously. While a bud-type reaction in calculi without domains would result in an extruded membrane that would also include  $\sigma_5$  and  $\sigma_6$ , bud with domains would cause the extrusion of the brane defined by the domain locally involved in the bud reaction.

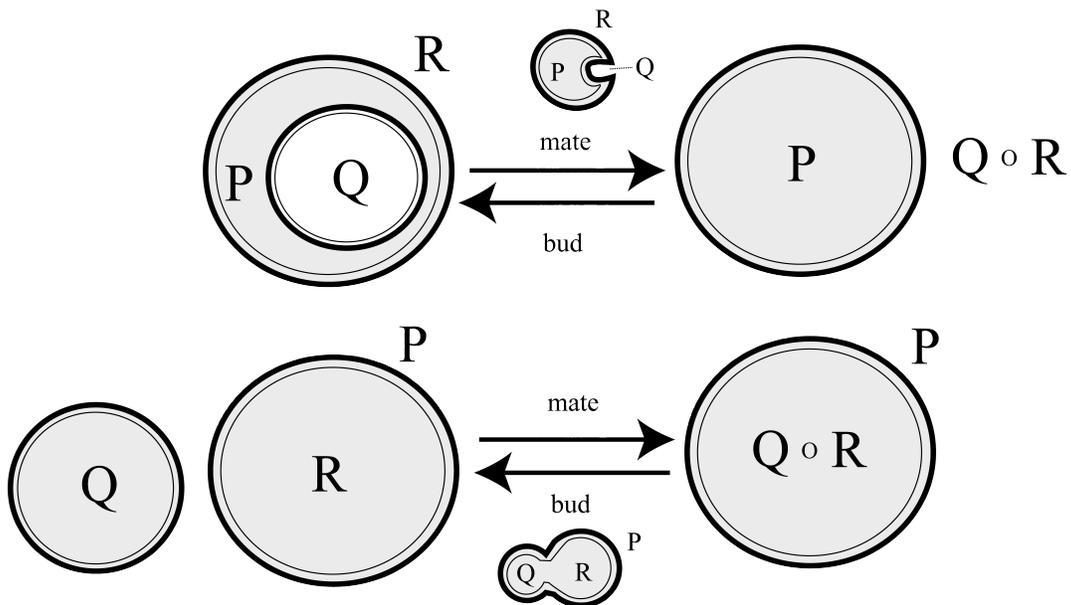


Figure 3.5: Graphical representation of PABM reduction rules. These rules generalize *endo*, *pino* and *exo* reactions (top) and *mate*, *bud* and *drip* reactions (bottom) as *mate* and *bud* type reactions.



## Chapter 4

# Modeling gold NP uptake and intracellular distribution

Creating a general model for NP uptake and intracellular distribution is an important complementary approach to experimental investigations in NP efficiency and cytotoxicity evaluation. For efficiency estimation, we are mostly interested in finding out how much time is needed for endosomal escape as a function of the physico-chemical features of the NP. In establishing the first model, we wanted to have an idea of how much of the particles localize in the nucleus, based on empirically measured aggregation, uptake, endosomal escape, cytosolic transport and nuclear entry rates, as well as known physical constraints provided by the cell. The model should, of course, capture experimentally-derived intracellular NP distribution before it is extended for cytotoxicity modeling, but extensibility should be an expected feature. Specifically, we need to at least incorporate details on how NPs interact with or influence the transcription of genetic material, perhaps going as far as identifying specific genes and gene products, and detailing the signaling pathways that influence its expression, for which experimental evidence exists [62, 41]. Compatibility with existing models of signaling and gene expression pathways, mostly written in SBML [63], is also a non-trivial considerations in modeling. There is a program that converts SBML models into SPiM which should allow for a possibility of integration with pathway models [64].

Most NP cytotoxicity models are quantitative structure-activity relationship (QSAR) models. These are essentially regression models that describe the relationship between biological activity and the physicochemical and structural characteristics of compounds of interest, or related compounds that have been characterized [65, 5]. It should be noted that QSAR may also be used to predict physicochemical properties themselves [4], although there is no current synthesis on how far this has been exploited and how successful it has been for NPs, which possess different physical properties from its source materials. It is rather clear, however, that the results from QSAR do not answer *why* an NP is cytotoxic or not, but rather *which* categories of NP characteristics are most likely to influence its toxicity; validation experiments are directed towards evaluation of the robustness of these categories [4]. Categories that have been proposed for “nano-QSAR” include size distribution, agglomeration state, shape, porosity, surface area, chemical composition,

structure-dependent electronic configuration, surface charge, and crystal structure [4, 7]. Executable biological models, on the other hand, are expected to yield testable mechanistic scenarios of how a particular NP causes cytotoxicity; in this sense, results from QSAR can be viewed as complementary, rather than competing. In this chapter, we discuss how we built our generic model using gold NP data, highlighting the advantages of creating it in pi calculus.

## 4.1 Choice of gold nanoparticles as a case study

One of our first considerations in setting up a generic model of NP uptake and intracellular movement is data availability, particularly with respect to kinetic parameters. This minimizes the need for parameter estimation, in turn increasing the probability that disparities between experimental and model-derived data result from the model structure itself, and not from wrong parameter estimates. Gold NPs are fairly well-characterized, given its wide array of existing and potential biological applications, including its use as contrast agents, drug and gene delivery vectors, and as a heat source in anti-cancer hyperthermia therapy [66]. This implies that data exist on its uptake [67, 68] and aggregation [69] characteristics, as well as its potential cytotoxicity [41, 70, 71]. These assays have involved both modified and unmodified gold NPs of different sizes, tested on different cell types [67].

## 4.2 Nanoparticle aggregation, uptake, and combinatorial explosion

The influence of NP size on uptake, which is effectively determined by NP aggregation characteristics is challenging to model using ODEs, especially if aggregates of different sizes have to be tracked explicitly during a simulation. This is foreseeably the case in toxicity [72, 73] and degradation-related problems [74], where size plays a major role. For instance, a 4 nm NP, as well as 4 nm NP clusters with diameters ranging from 15 - 20 nm, appear to enter nucleus easily Paper C, [70, 71], where it could interact with the DNA, and cause cytotoxicity [75]. The moment, however, that it forms an aggregate with a diameter of between 9 [33] and 26 nm [76], which are the limits for passive diffusion and signal-mediated transport into the nucleus, respectively, then nuclear entry would no longer be physically possible, unless mitosis occurs [18]. In order to capture these in an ODE model it is necessary to define  $m * n$  equations, each describing the evolution of  $1..n$  aggregate sizes through the  $m$  states or locations that these could take. The aggregation reaction itself, however, is a reaction that can be generalized as  $NP(\text{size}_1) + NP(\text{size}_2) \rightarrow NP(\text{size}_1 + \text{size}_2)$ , with the base case being  $NP(1)$  – that is, an unaggregated NP. This possibility of generalization makes the pi calculus model significantly less cumbersome. It also permits easier expansion to include other reactions involving aggregates of a particular size. It is noteworthy to mention at this point that the way that processes are implemented in SPiM precludes the possibility of an instantiation of the  $NP$  process from reacting with itself.

### 4.3 Pi calculus model of gold NP uptake and intracellular distribution

In Paper C, we describe a pi calculus model for nanoparticle aggregation, size-restricted uptake and movement within the cellular compartments. Briefly, we defined the following rules, detailing inter-NP interaction and NP-cell interaction; the rules are written as reaction-type statements for readability:

1. Aggregation:  $\text{NP}(\text{size}_1) + \text{NP}(\text{size}_2) \rightarrow \text{NP}(\text{size}_1 + \text{size}_2)$
2. NP-cell Binding and internalization:  $\text{NP}(\text{size}) + \text{Cell} \rightarrow \text{NP}(\text{size})\text{-Cell} \rightarrow \text{Cell} + \text{Endosome}(\text{NP}(\text{size}))$
3. Endosome-Endosome fusion:  $\text{Endosome}(\text{NP}(\text{size}_1)) + \text{Endosome}(\text{NP}(\text{size}_2)) \rightarrow \text{Endosome}(\text{NP}(\text{size}_1), \text{NP}(\text{size}_2))$
4. Transcytosis:  $\text{Endosome}(\text{NP}(\text{size})) + \text{Basolateral membrane} \rightarrow \text{Basolateral membrane}^1$
5. Endosome lysis:  $\text{Endosome}(\text{NP}(\text{size})) \rightarrow \text{Cytosol}(\text{NP}(\text{size}))$
6. Nuclear entry:  $\text{Cytosol}(\text{NP}(\text{size})) \rightarrow \text{Nucleus}(\text{NP}(\text{size}))$

Cell compartments were treated as processes that can send and receive the number of NPs or NP aggregates that enter it. The initial runs were performed without incorporating the effect of aggregate size on the uptake rate. The main results from the model are in the form of spatio-temporal distributions of NPs (Figure 4.1). The model results, however, feature generally higher nuclear translocation of NPs. In order to improve the data fit, as well as to demonstrate the ease by which one could implement changes in a pi model, we introduced a simple restriction of optimal uptake rates for aggregates of size 43 - 125 NPs. Although uptake is several steps upstream of nuclear entry, this modification already has a large impact on predicted nuclear entry, as well as the number of clusters that can be found in the nucleus.

What this work primarily demonstrates is how rule-based models can be used in understanding delivery information better. Using this base model, it becomes possible to ask questions as ‘What is the effect of having  $n$  cytosolic NPs on a signaling pathway  $P$ , considering that NPs release ions that react with proteins  $X$  and  $Y$  in  $P$ ?’ Similarly, we could use the model to ask to what extent does inter-NP variability have an effect on NP efficiency – a question that we answer in the next chapter.

---

<sup>1</sup>Transcytosed NPs are assumed to be unable to reenter the cell, and are treated as lost

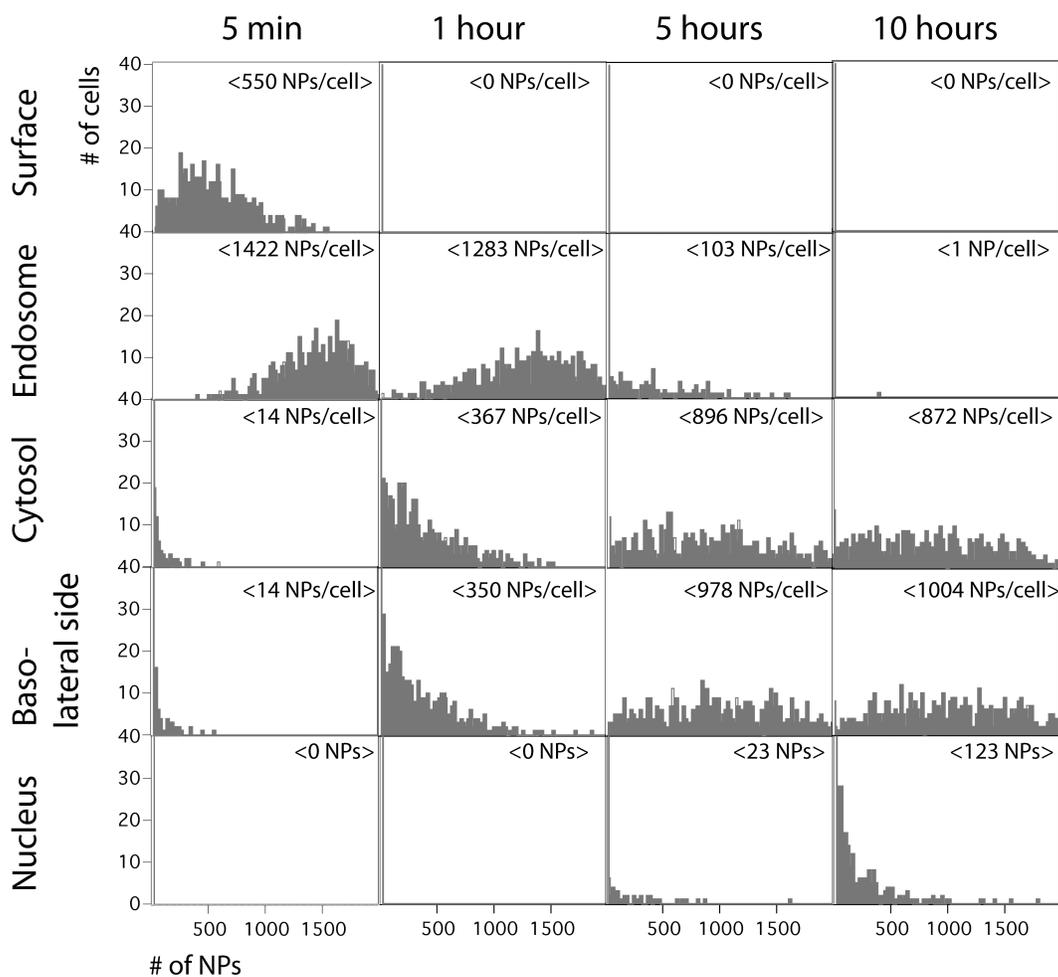


Figure 4.1: Predicted spatio-temporal distribution of gold NPs across 500 cells. Model results are generally consistent with reported NP intracellular distribution up to five hours; the average number of nuclear NPs at 10 hours, however, is higher than expected. We subsequently reduced this number by imposing a limit to the aggregate sizes that are taken up at an optimal rate, based on known size effects in literature [68]. Figure from Paper C.

## Chapter 5

# Experiments and modeling of endosomal escape events

Endosomal escape has been identified as the main bottleneck in gene and drug delivery [27, 19]; it has been consequently of interest to develop both experimental techniques and complementary modeling techniques that would allow us to analyze endosomal escape characteristics of different vectors, including the elucidation of the *mechanisms* that permit some vectors to perform better than others. There are, however, few real-time, single endosome lysis studies because of difficulties in detecting the events reliably. Previous studies of vector or viral tracking involved labeling of multiple cellular and vector/viral components, and used confocal laser scanning microscopy was used [77, 78, 79, 80, 81]. In Akita et al., at least 20 z-slice images for each of 30 cells were reported to yield statistically meaningful tracking results [77]. If this method were applied to a 30-hour transfection experiment [18], this would entail the capture, consolidation and analysis of 216,000 images. Due to these technical limitations, studies have mostly concentrated on spatial distribution of the material, without the temporal aspect [79, 80]. Endosome escape efficiency is only inferred from a minimal amount of sampling points [77]. Having a direct method for analyzing uptake and endosomal escape will not only yield time distributions of the escape events, but also information on *other* factors that might influence the time distribution such as the vector load per endosome and cell-specific endosome size distributions. Such a method will be powerful in exploring the performance of a delivery vector, especially if information is obtained for different cell types.

To accomplish this, we needed a reporter system that could either be co-delivered with other vectors, or be used as a vector itself. The ideal system should have different signals for each target component, and should not diffuse so easily in the cytosol. We chose the recently characterized colloidal mesoporous silica (CMS) particles, developed in the lab of Prof. T. Bein, as the reporter system core, to which lytic agents such as PpIX and dyes such as ATTO633 covalently linked to a quencher QSY21 can be attached [82]. PpIX is a photolytic agent that produces reactive oxygen species (ROS) on activation, which oxidize double bonds in the both the vector membrane and the endosome membrane. On vector movement to the cytosolic milieu, the bond between ATTO633-QSY21 is reduced,

resulting in fluorescence that is localized to the NP. As with our previous work, we need to complement the data with a model from which mechanistic assumptions can be derived. In this chapter, we discuss experiments performed using the CMS reporter system, as well as a mechanistic model that we have generated in order to describe all the data obtained.

## 5.1 Characterization of endosomal release of a lipid-coated nanoparticle reporter system

### 5.1.1 Generation of lipid-coated nanoparticles

To deposit a lipid bilayer on the CMS particles, we developed a solvent-exchange based protocol that gradually moves from amphiphatic ethanol solution to pure aqueous solution (Paper D). Some previous protocols have either used a film-hydration method [83] or a similar principle of solvent exchange [84], but are limited either by the difficulty of the protocol, or the incompatibility of the final solvent with biological samples. Our protocol directly uses dehydrated lipid that is resuspended from ethanol to water. We have shown that the protocol generates fluid liquid bilayers on a flat surface as shown from a continuous bleaching protocol (Paper D, [85]). Additionally, correlation experiments show that the method yields intact bilayers on CMS particles for several lipid compositions [86]. A lipid coat deposited in this manner also provides better stability against CMS NP aggregation in aqueous medium [86].

### 5.1.2 Reporter system characterization and experiments

The reporter system in the lipid-covered form has been extensively characterized in the work of A. Sauer from the group of Prof. C. Brauchle, as well as by H. Engelke, with the membrane quality being assessed using confocal microscopy and correlation spectroscopy [86]. Here, as well as in Manuscript F, we provide further characterization of the system using single-cell uptake experiments and single-endosome lysis experiments.

As a proof-of-concept experiment, we used a release-on-demand system PpIX, for which the endosome lysis time distribution is expected to be narrow. A typical experiment to obtain the endosome lysis times is shown in Figure 5.1. Briefly, cells are incubated with PpIX-decorated ATTO633-QSY21 NPs and ATTO633-QSY21 NPs negative controls for 12 hours. Uninternalized antibodies were washed off and excess fluorescence is quenched prior to PpIX excitation with a 405 nm laser. Excitation results in the production of reactive oxygen species (ROS). ROS-mediated oxidation of double bonds in lipid tails destabilize both the NP and endosome membrane. Experiments were performed on a fibroblast cell line, which is expected to take up a fairly high amount of NPs, as opposed to the carcinoma cell line against which it was tested. Apart from cell-related effects, we also tested the effect of the lipid membrane composition, as well as the duration of activation. Fluorescence images were taken every second over two minutes at on three emission channels. Endosomes were semi-automatically detected using an in-house ImageJ plugin [87]; undetected

endosomes are segmented manually Figure 5.1. Since PpIX activation halts the movement of endosomes, it is only at the first frame that segmentation is performed.

**Statistics from single-endosome PpIX-mediated lysis with different lipids, cell types, and PpIX activation times**

Time distributions of DOPE@CMS exhibit marginally earlier lysis times than DOPC@CMS, indicating that kinetically, the fusogenic effects of DOPE are superseded by PpIX. Our results also suggest that the activation time appears to play a more prominent role in Renca-LacZ than in 3T3, as evinced by the bigger shift in the distribution when the activation time is decreased (Figures 5.2C and 5.2D). Furthermore, the distribution of lysis times after a two-minute PpIX activation is considerably more spread than in 3T3, which could either be an experimental artefact, or an indication of cell-specific dependence of endosome lysis.

To check if there were cell-specific differences that influenced lysis timing, we obtained the endosome size distribution based on the regions of interest in the AlexaFluor Dextran 488 channel. We also obtained the corresponding intensity at the red channel as a rough approximation of the endosome load. The disparities in the endosome area (Figure 5.3A) and endosome load (Figure 5.3B) are relatively marked between the two cell types, with Renca-LacZ cells displaying more size variation, and tending towards larger endosomes. This could be due to a higher expression and/or a longer lifetime of Rab5 in Renca-LacZ, a protein that controls endosome fusion, and consequently endosome size distribution [88]. The observations are consistent with a recent theoretical study that shows that the endosomal escape time increases with the endosome size, and decreases as the net amount of membrane disruptive material increases [21].

**Cell-type dependence of nanoparticle uptake**

Due to the unexpected disparities in the lysis results, we wished to quantify the uptake differences between 3T3 and Renca-LacZ, as well as two other cell lines that we have not previously used. These include Beas2B, a normal human bronchial epithelial cell line; and Huh7, a hepatocarcinoma cell line. Parallel uptake experiments were performed on the cells, which were incubated with NPs at two hour intervals up to eight hours. Our results confirm that 3T3 cells take up more DOPC than Renca-LacZ cells on average. Additionally, the cell-to-cell variability of uptake, as evinced by the distribution width, is approximately twice as high for Renca-LacZ than 3T3 (Table 5.1).

Cell type uptake dependence of non-functionalized NPs has been reported previously [89]. There have also been recent reports on uptake differences resulting not only from cell surface properties, but also cell type-specific cytoplasmic and nuclear pore penetration constraints [90]. These results altogether show the need for investigating this phenomenon more closely, especially in the context of NP design and dosage.

| Sample               | Mean              | Width              |
|----------------------|-------------------|--------------------|
| DOPC@CMS, 3T3        | $675.7 \pm 0.3$   | $17.2 \pm 0.2$     |
| DOPC@CMS, Renca-LacZ | $644.5 \pm 1.7$   | $41.7 \pm 2.4$     |
| DOPC@CMS, Huh7       | $673.9 \pm 1.2$   | $20.9 \pm 1.1$     |
| DOPC@CMS, Beas2B     | $726.9 \pm 9.4$   | $132.5 \pm 14.3$   |
| Latex, 3T3           | $1314.1 \pm 29.9$ | $227.8 \pm 43.3$   |
| Latex, Renca-LacZ    | $1603.5 \pm 19.3$ | $1065.7 \pm 31.6$  |
| Latex, Huh7          | $613.4 \pm 69.4$  | $1314.5 \pm 102.0$ |

Table 5.1: Selected gauss fit parameters of cell-type specific NP uptake profiles at eight hours post-incubation

## 5.2 Stochastic pi calculus model of PpIX-induced endosomal escape

To model the endosome lysis data, we extended our gold NP model to take cell-dependent uptake, NP variability in terms of PpIX load per particle, and the PpIX-mediated endosome lysis reaction into account. The structure of the model is shown in Figure 5.4. For our runs, we made the initial assumptions that the rates from the gold NP model hold; given that the CMS particles are 100 nm, while the gold NPs for which the model was written are 4 nm, we changed the limits of optimal uptake from 43 - 125 NPs per aggregate to a maximum of 15 per aggregate. We also preceded the uptake process with a delay that generically represents cell-specific uptake differences. Each of 10000 NPs was initially assigned a different PpIX load based on a normal distribution centered at 50 PpIX/NP and a width of 5.

For the first version of the model, we omit effects brought about by the choice of lipid. As a base case for the fit, we consider the data for DOPC-covered particles in 3T3 cells. Simulations indicate sensitivity to endosome-specific parameters, specifically, the number of net PpIX per endosome required to cause endosome lysis (Figure 5.5). However, changing this parameter is not expected to be sufficient to capture the behavior in Renca cells, for which the influence of heterogenous endosome size on this minimum number probably requires explicit representation. The current model is limited by the fact that it is not possible to use a random number generator within a running SPiM code; this would be necessary for assigning a size attribute to the dynamically created endosome process, which we infer would have an influence on the lysis time [21]. This feature is currently pending as a request for integration into SPiM.

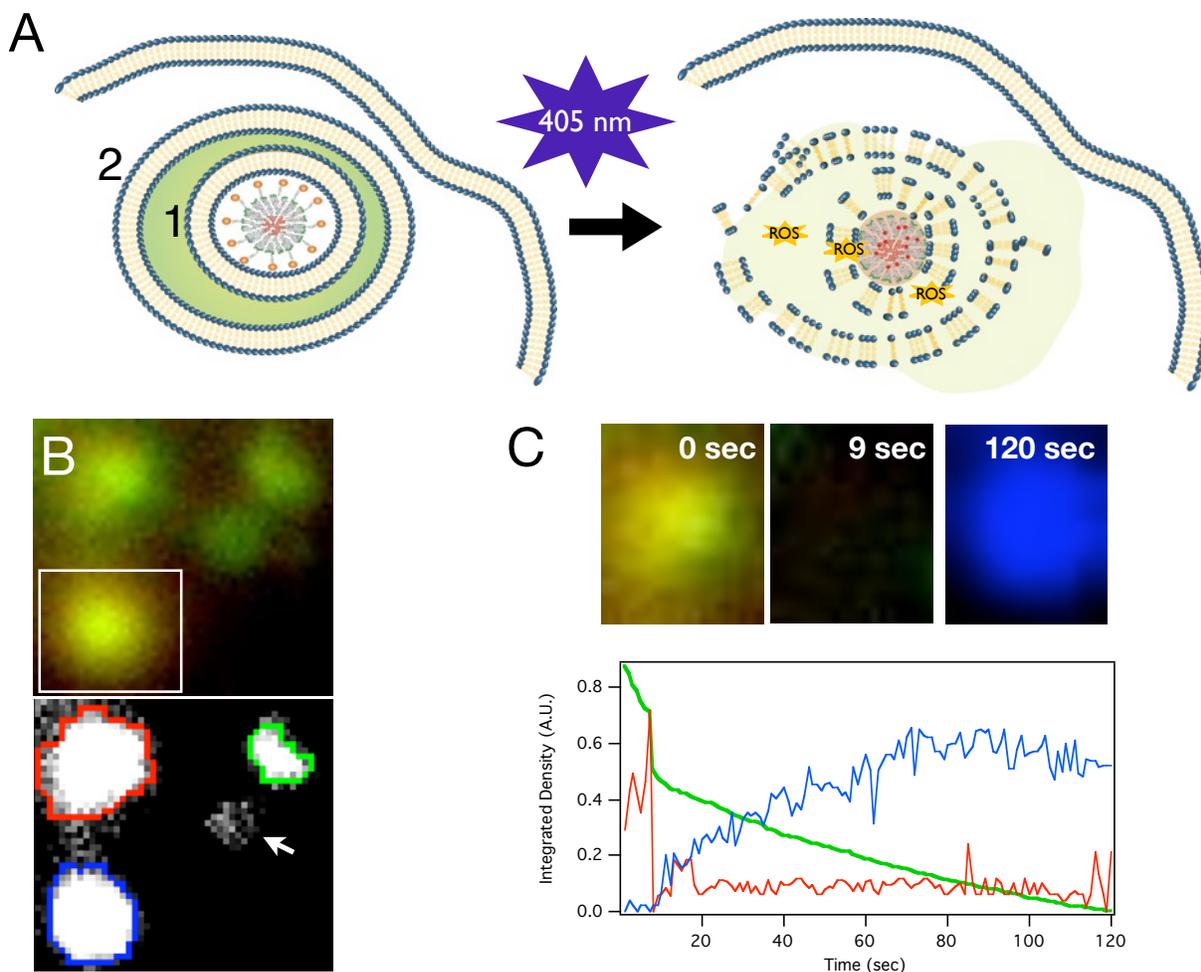


Figure 5.1: Our reporter system is an NP comprised of a colloidal mesoporous silica (CMS) core to which the photosensitizer protoporphyrin IX (PpIX), as well as ATTO633 quenched by covalently-attached QSY21 (A1). The system is co-introduced with a fluid-phase marker, Alexa-Fluor Dextran 488, which delineates the boundaries of the endosome (A2) The whole NP is sealed by a lipid bilayer. On activation by a 405 nm laser, reactive oxygen species are produced and oxidize double bonds in the lipid tails, resulting in membrane destabilization. The corresponding microscopy images prior to (B) and during (C) lysis are shown. The corresponding segmented images in (B) are shown; not all endosomes are successfully detected, and these have to be manually defined (arrow). A typical signature of endosome lysis is shown in C, where a drastic drop in the red and green channels, and a gradual rise in the ATTO channel, is detected. Automatically detected discontinuities, indicative of lysis are recorded for different lipid coats and different cell types.

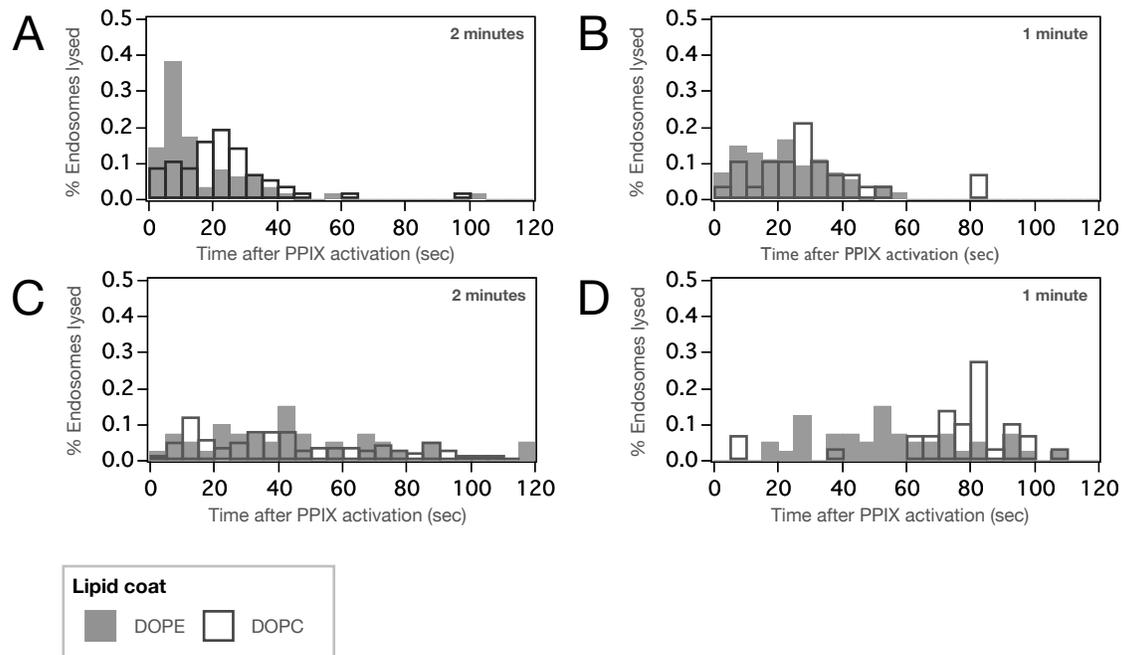


Figure 5.2: Time distributions of endosome lysis events after PpIX activation in 3T3 (A and B) and Renca-LacZ (C and D) cells. PpIX-functionalized DOPC@CMS and DOPE@CMS were incubated overnight, then activated for either one or two minutes. Results generally show marginally earlier lysis times for DOPE@CMS. It is interesting to note that lysis appears to be more concentrated towards earlier times for 3T3 cells than for RencaLacZ; furthermore, the shift in the lysis times when activation is shorter is less prominent in 3T3.

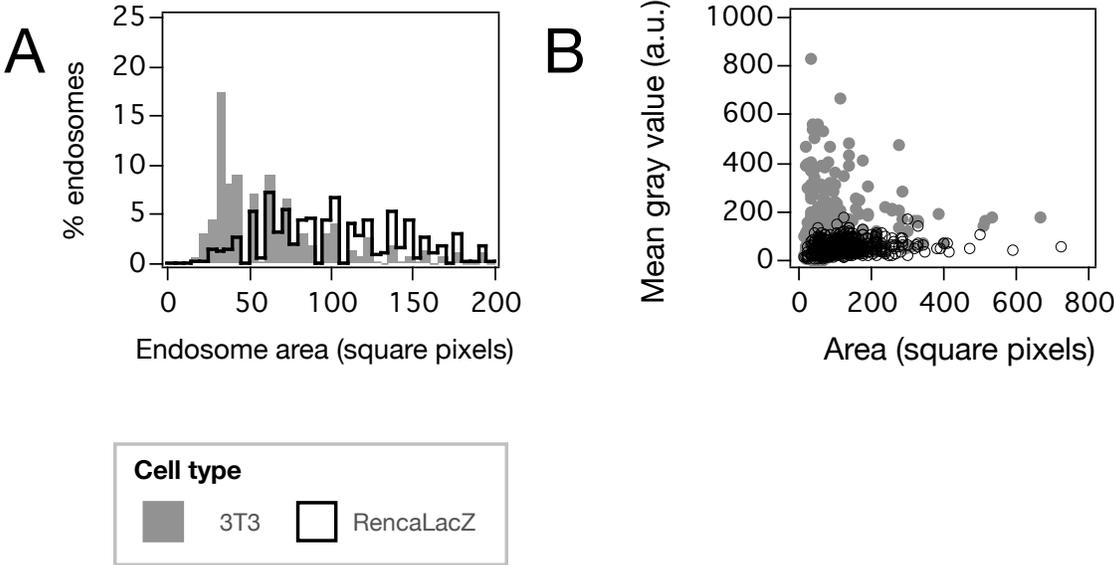


Figure 5.3: Endosome size distribution (A) and NP colocalization intensities per unit area in an endosome (B) in 3T3 and Renca-LacZ cells.

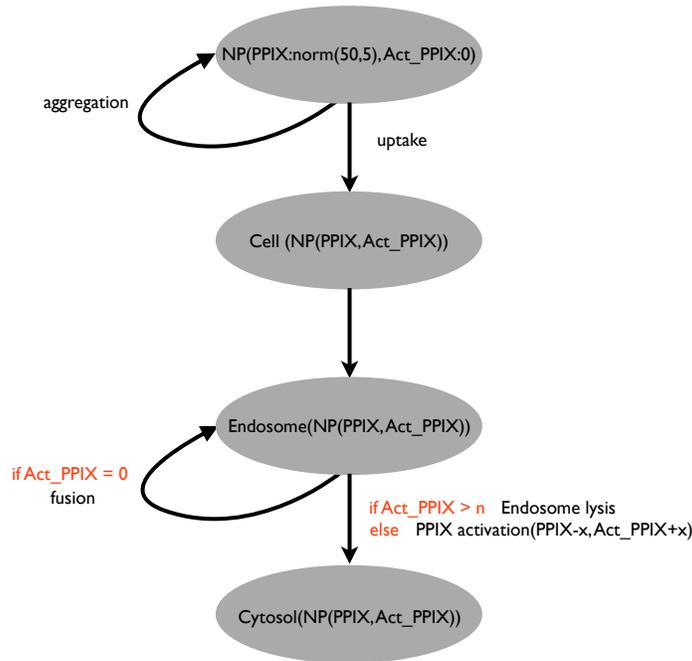


Figure 5.4: SPiM model structure for PpIX-mediated endosomal escape. Each NP process is initially assigned a PpIX load attribute from a normal distribution centered at 50 with a width of 5, as well as a parameter representing the number of activated PpIX, Act\_PpIX. NPs can bind to a cell or form aggregates. Cell binding can trigger uptake, whose rate is adjusted to reflect differences in cell type. Each uptake event results in the generation of an endosome; endosome fusion can occur as long as Act\_PpIX = 0. On activation, a subset of the net PpIX associated with the endosome is activated. If this is greater than a threshold  $n$ , then lysis occurs. Otherwise, activation continues until  $n$  is reached.

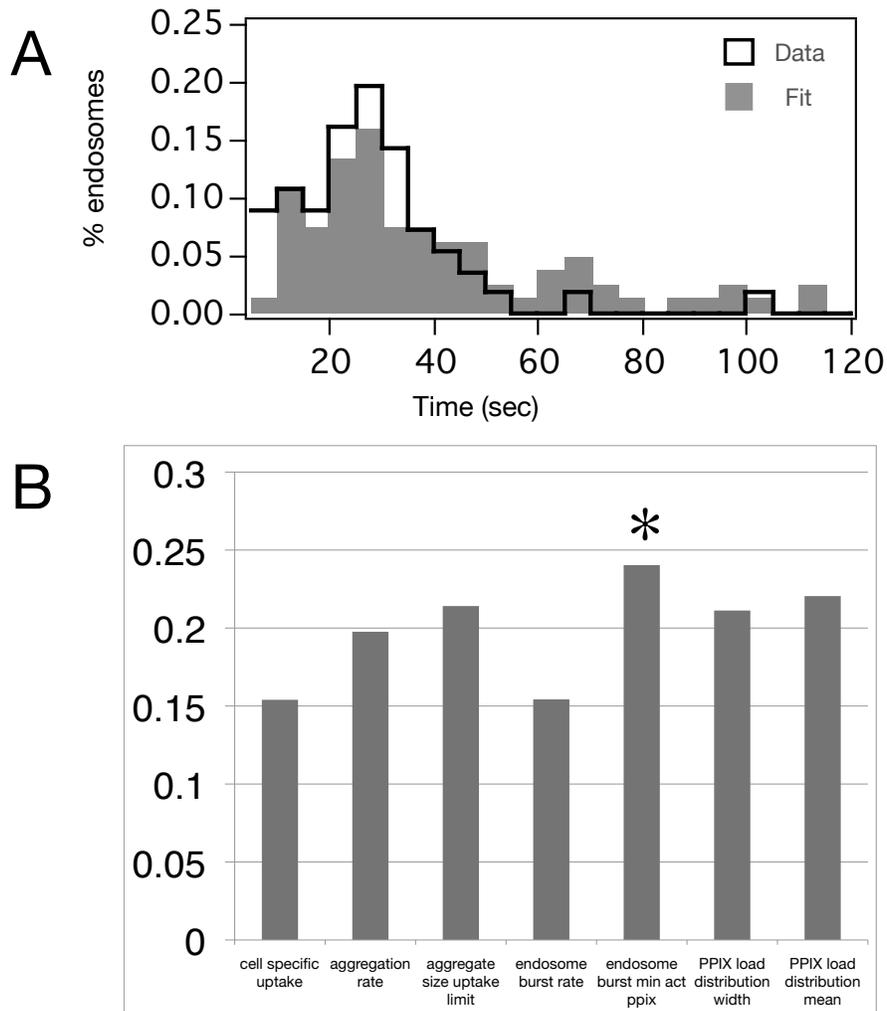


Figure 5.5: SPiM model fit of lysis data for DOPC-covered particles in 3T3 cells (A) and the corresponding parameter sensitivity analysis (B). The model is most sensitive to changes in the minimum number of PpIX per endosome (marked by an asterisk). Variations in NP batches have the potential to affect the results significantly as well, but our fit indicates that inter-NP variation is less than 10%.



## Chapter 6

# Design considerations for nanoparticles with protein or peptide modifications for targeted delivery and controlled release

The previous chapter has dealt with a release-on-demand system that is suitable for delivery of cargo to tissues for which light penetrance is not an issue. For other tissues, it would be interesting to develop NPs that do not suffer from this limitation. As indicated in Chapter 2, viruses have been the main source of inspiration for such modified NPs. Current hybrid vector designs frequently include targeting and lytic peptides, rather than full proteins, to improve uptake and endosomal escape, while minimizing the likelihood of an immune response [25]. These peptides sometimes bear mutations that are known to enhance its function. Peptides, nonetheless, can still be immunogenic, and some work has been done both in terms of defining the minimum peptide length that still results in fusion [91, 92]. In comparison with the full protein, however, the performance of peptides is often not as efficient. For instance, the transduction domain of the TAT protein of HIV-1, which is sufficient to mediate viral material transfer into the cell in its native protein form [93, 94], needs to be coupled with another fusogenic peptide from influenza hemagglutinin (HA) to escape micropinosomes [95]. Similarly, longer cell penetrating peptides have been reported to generally display higher uptake than shorter counterparts, even if uptake is receptor independent [96]. Frequently, the structural features of a peptide, which are necessary for fusion, are omitted in the use of shorter peptides, or peptides with modified sequences [97]. Finally, larger-scale structures such as the fusion pore also have to be taken into account in hybrid vector design. Studying the mechanisms of action of a native targeting or fusion protein is consequently a core feature in this area.

## 6.1 Modeling minimal requirements for HA-mediated endosomal release

Although it is not currently possible to precisely control the number of engineer surface functionalizations or preassembled fusion pore structures on NPs, the incorporation of a minimal number of targeting ligand is of particular interest, since high ligand densities could lead to non-specific interactions with non-target tissue, as well as increased immunogenicity [25, 98]. As a starting point, we started looking at the mechanistic details of HA-mediated endosomal release, specifically the minimal number of HA molecules required to form a fusion pore. Perhaps not so surprisingly, the reports that we found on these are contradictory at first glance [99]. To understand these results better, we compared the experimental methods, data analysis and experimental results; in the six studies considered, we found that the experiments could generally be classified into cell-cell and virus-cell fusion setups, which expectedly yield results that are not directly comparable. Limiting our analysis to virus-cell fusion setups showed that the data are almost the same. The interpretations differed based on the phenomenological model used. There were also cases when different results are interpreted to yield the same minimum mechanistic requirements for fusion because different phenomenological models were used to analyze them; of course, this implies that using one model to analyze the other set of data would have resulted in conflicting, rather than converging conclusions. To resolve these inconsistencies, we created a rule-based model, detailed in Paper E, that was first specified in PABM, then executed using the model-checker, PRISM [100].

The different possible combinations for the pore size, which ranges from six to eight; and number of fusogenic HAs in a pore, which was reported to be between one and three, makes it practical to first reduce the solution space and limit the model testing to the most feasible scenarios. To achieve this, we first ran simulations using the fastest reaction possible for pore sizes ranging from five to nine. The closest to the data was a pore size of six, which we then used for testing all the possible combinations of fusogenic and non-fusogenic HAs. We obtained the best fit using a fusion pore size of six, of which three are fusogenic; of these three, two are free HA, and one is bound. These results were subsequently verified by using the same rates obtained from virus-cell fusion data fits, together with HA and sialic acid (SA) density estimates in the cell-cell fusion scenario to fit cell-cell fusion data [99]. These results may eventually be used in designing NPs with preassembled fusion pore-forming units, which would not only allow the NPs to escape from the endosome, but possibly to enter the cell by direct fusion. Again, if it becomes possible to control the number of surface functionalizations in an NP, it would be interesting to test uptake and fusion kinetics using such minimally-modified vectors.

## 6.2 Outlook

### 6.2.1 NP uptake insights

Our understanding of interactions between NPs and cells, as well as the effect of serum protein association with NPs, particularly those that form the corona, on cell-specific responses, is probably just at the surface of the knowledge that is necessary to design NPs with more predictable behavior. For instance, the experiment and model results in Chapter 5 show how much cell-type specific characteristics, and not inter-NP variations such as PPIX load, define the spread of endosome lysis events. However, we can only speculate on why the uptake is varied, and our data and modeling results indicate the possible involvement of Rab5 expression levels and lifetime as a *contributor* to the difference between CMS performance in Renca-LacZ and 3T3 cells. Nonetheless, this does not tell us *why* uptake among the different cell types is different.

We next asked the question of how much CMS NP uptake patterns will change by functionalization. In a proof-of-concept experiment, we took filamentous HA from *B. pertussis*, a heparin binding protein which induces lipid destabilization and mixing [101, 102], and resuspended it with the DOPC cover of the CMS NPs [103], effectively co-introducing it with the NPs. We chose the protein because of its targeting and destabilizing properties, its possible structural similarities to other fusogenic peptides such as influenza HA and integrins [104], as well as its commercial availability. These uptake profiles could also be considered as an indirect measure of the availability of the binding partner in the target cell, and could be used as a starting point for evaluating the targeting peptides that would work better for specific cell types. Uptake results without (Figure 6.1A) and with (Figure 6.1B) HA are shown for the same cell lines tested previously. Interestingly, fibroblast cells still show the least cell-cell variability in terms of uptake, but not the highest average NP load, which is now associated with Huh7. There is, however, a general uptake increase and spread for all cells, giving us again an idea that uptake may be generally improved with the use of targeting proteins, but would significantly magnify the uptake response variability.

In future work, it would be interesting to analyze the performance of arrays of NPs with different functionalizations, deposited in a controlled manner, against different cell lines. A study where an NP library, comprised of NPs modified with 146 small chemical groups, was constructed and tested on five cell lines, may serve as a template for further assays [105]. The information that can be derived from such an experiment include information on surface modifications that change the affinity of particles; material that discriminate between functional states of a cell and *ad hoc* identification of disease-specific agents without knowledge of the target [105]. If the study is modified to use peptide, rather than small molecule functionalizations, the study would also yield information on surface protein expression profiles across different cell types; this is, by itself, is an important result, given that information is fairly scarce with respect to protein profiles for cell lines. The use of the NP library could be extended to characterize the protein coronas for different NPs. It would be interesting to see how much uptake characteristics are affected by the

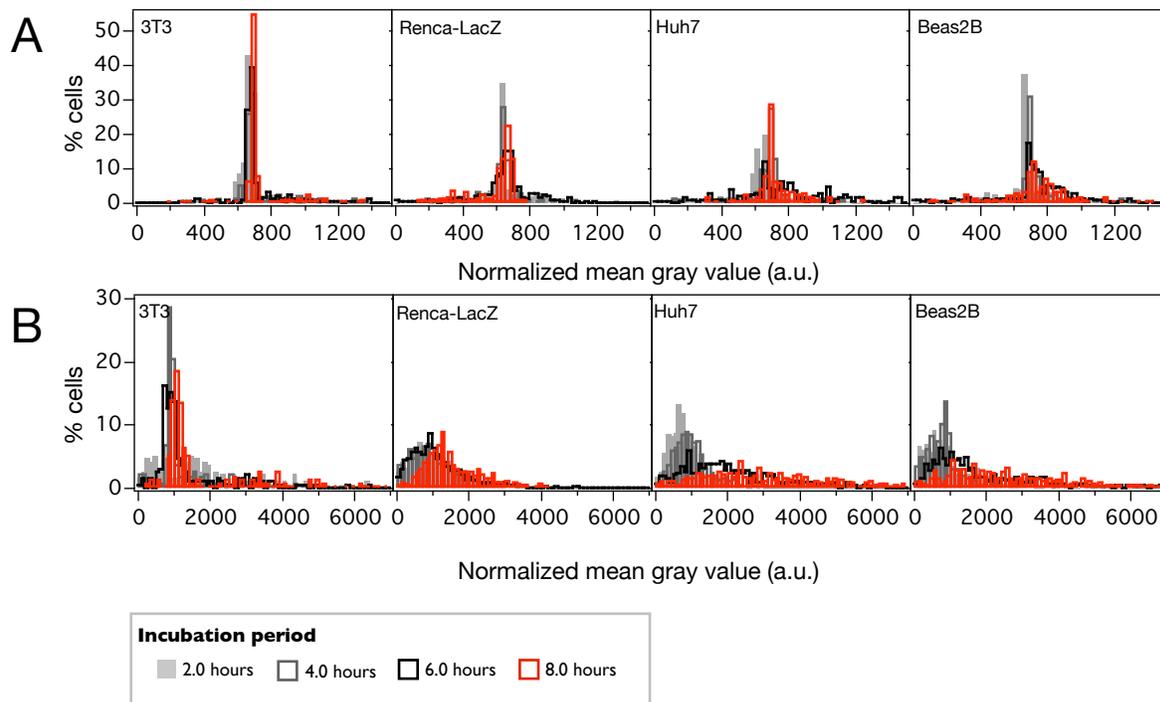


Figure 6.1: Uptake of DOPC@CMS NPs in 3T3, Renca-LacZ, Huh7 and Beas2B cells in the absence (A) or presence (B) of filamentous HA from *B. pertusis*. Note that even with non-specific insertion into the membrane, there is at least a two-fold increase in uptake. The larger spread of data could be partly expected, both as a result of the variation in the cell surface expression of heparin across the different cell types, as well as the varied quantities of HA inserted per membrane.

corona, and how NP functionalization affects the corona itself.

### 6.2.2 Future work in nanoparticle design

The successful design and testing of a protocell, which combines the properties of liposomes and nanoporous NPs, was recently reported [25]. This protocell was comprised of a lipid-covered CMS NP to which a modified influenza HA lytic peptide and a human hepatocellular carcinoma targeting peptide have been incorporated. The targeting peptide makes uptake highly specific and more efficient, with an increase of internalized protocells by an order of  $10^2$  in target cells. Uptake saturation is observed at  $t = 1$  hour post-incubation, which is five hours earlier than what we have generally observed for unmodified particles. Interestingly, DOPC protocells show improved surface binding to its target cells independent of peptide density, which has been reported to range from 6 - 2048 peptides per particle. This is attributed to the enhanced bilayer fluidity, which permits peptide recruitment to the binding site. A problem with this, however, is that the extremely high

membrane fluidity results in cargo leakage; stable formulations for drugs require bilayers comprised of saturated lipids such as DSPC, and high cholesterol concentrations [25, 106]. For such less fluid formulations, more peptide would be required to generate equivalent uptake effects [25]. Since the endosomal escape of HA-modified DSPC protocells was not covered in the study, it could be speculated that this would be significantly lower than for DOPC protocells, if not completely absent, especially for lower HA concentrations. This is because of the requirement for membrane fluidity to assemble the fusion pore-forming unit [99].

In cases such as this, the use of pre-assembled pore-forming complexes might be an interesting option (Figure 6.2). While artificial pores that have been created so far are limited to lipid nanotubes [107], the creation of preassembled protein or peptide pores using DNA origami [108, 109, 110] might be a viable option. It is in cases as these where combining results obtained from modeling, as in Paper E, and experimental techniques, might prove practically beneficial.

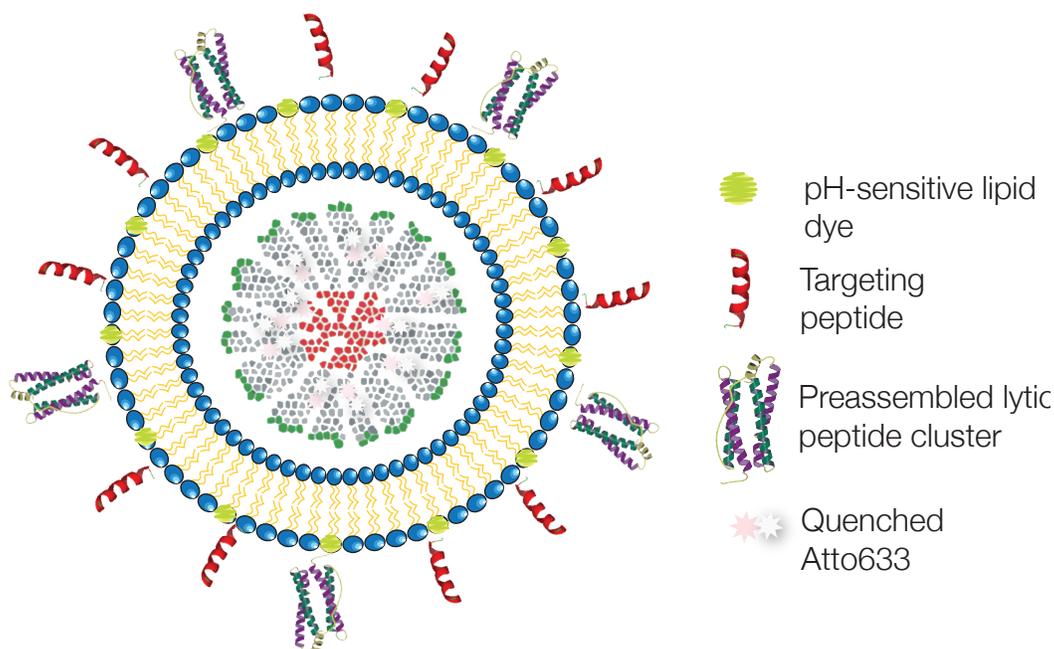


Figure 6.2: Design for a CMS-based NP tracker and delivery system with pre-assembled pore-forming complexes. This inclusion of these pre-assembled pore-forming complexes would facilitate endosomal escape even if the encapsulating bilayer is not fluid, as the case would be when lipids such as DSPC are used. Less fluid bilayers are often recommended for drug carriers, as these prevent premature cargo leakage [25]. Note that the design also incorporates a pH-sensitive lipid dye, possibly enabling real-time tracking from incubation, rather than after incubation. The targeting peptides might be dispensed with, if these are pre-assembled with the pore-forming cluster as well.



# Appendix A

## Original publication P1

Gerlinde Schwake, Simon Youssef, Jan-Timm Kuhr, Sebastian Gude, Maria Pamela David,  
Eduardo Mendoza, Erwin Frey, Joachim O. Rädler

**Predictive Modeling of non-viral Gene Transfer**

*Biotechnology and Bioengineering*, **105**: 805 - 813, 2010.



# Predictive Modeling of Non-Viral Gene Transfer

Gerlinde Schwake,<sup>1</sup> Simon Youssef,<sup>1,2</sup> Jan-Timm Kuhr,<sup>1,2,3</sup> Sebastian Gude,<sup>1</sup>  
Maria Pamela David,<sup>1</sup> Eduardo Mendoza,<sup>1,2,4</sup> Erwin Frey,<sup>1,2,3</sup> Joachim O. Rädler<sup>1,2</sup>

<sup>1</sup>Fakultät für Physik, Ludwig-Maximilians-Universität, Geschwister-Scholl-Platz 1,  
D-80539 München, Germany; telephone: +49-(0)89-2180-2438; fax: +49-(0)89-2180-3182;  
e-mail: gerlinde.schwake@physik.lmu.de

<sup>2</sup>Center for NanoScience (CeNS), Geschwister-Scholl-Platz 1, D-80539 München, Germany

<sup>3</sup>Arnold Sommerfeld Center for Theoretical Physics, Geschwister-Scholl-Platz 1,  
D-80539 München, Germany

<sup>4</sup>Institute of Mathematics, University of the Philippines, Diliman,  
Quezon City 1101, Philippines

Received 16 June 2009; revision received 17 September 2009; accepted 21 October 2009

Published online 1 December 2009 in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/bit.22604

**ABSTRACT:** In non-viral gene delivery, the variance of transgenic expression stems from the low number of plasmids successfully transferred. Here, we experimentally determine Lipofectamine- and PEI-mediated exogenous gene expression distributions from single cell time-lapse analysis. Broad Poisson-like distributions of steady state expression are observed for both transfection agents, when used with synchronized cell lines. At the same time, cotransfection analysis with YFP- and CFP-coding plasmids shows that multiple plasmids are simultaneously expressed, suggesting that plasmids are delivered in correlated units (complexes). We present a mathematical model of transfection, where a stochastic, two-step process is assumed, with the first being the low-probability entry step of complexes into the nucleus, followed by the subsequent release and activation of a small number of plasmids from a delivered complex. This conceptually simple model consistently predicts the observed fraction of transfected cells, the cotransfection ratio and the expression level distribution. It yields the number of efficient plasmids per complex and elucidates the origin of the associated noise, consequently providing a platform for evaluating and improving non-viral vectors.

*Biotechnol. Bioeng.* 2010;105: 805–813.

© 2009 Wiley Periodicals, Inc.

**KEYWORDS:** mathematical modeling; gene transfer; single cell; transfection/gene expression

## Introduction

Non-viral gene delivery systems have evolved over the last decade into widely used vectors for exogenous DNA delivery to eukaryotic cells. Synthetic cationic lipids and polymers, in particular, are used in molecular biology for transgene expression, and are being further refined for use in DNA-based therapies (Ferber, 2001; Patil et al., 2005; Roth and Sundaram, 2004). Despite considerable progress in the efficiency and characterization of vectors, important aspects of the delivery pathway and transfer kinetics remain poorly understood, including how artificial vectors are taken up, transported to the nucleus, and how these factors collectively influence the expression characteristics of a cell population. Current understanding from intracellular studies of transgene delivery includes the following steps: DNA-vector complex uptake via the endosomal pathway, followed by endosomal escape and cytoplasmic transport, nuclear entry, vector unpacking and transcription initiation (de Bruin et al., 2007; Kircheis et al., 2001; Lechardeur et al., 2005; Roth and Sundaram, 2004; Safinya, 2001; Suh et al., 2003). These processes are accompanied by a huge loss of material and temporal delays. It is therefore not surprising that transfected cells in a culture respond very heterogeneously over time, notably in terms of the expression onset time ( $t_{\text{on}}$ ) and the maximum expression levels attained. It is generally accepted that the expression behavior of a single transfected cell is stochastic, yet cell culture averaged expression levels are reliable indicators of gene transfer efficiency.

Flow cytometry is commonly used to measure fluorescence distributions over a population at a rate of up to 10,000 cells/second (Longo and Hasty, 2006). High-content single cell assays, in contrast, are particularly

Gerlinde Schwake, Simon Youssef, and Jan-Timm Kuhr contributed equally to this work.

Correspondence to: J.O. Rädler

Contract grant sponsor: Deutsche Forschungsgemeinschaft

Contract grant number: SFB486-B10; SFB TR12

Additional Supporting Information may be found in the online version of this article.

suitable for investigating the dynamics and heterogeneity of clonal cell populations, since individual cells can be followed with a high temporal resolution. In addition, quantitative image analysis has been successfully improved to reliably convert fluorescence intensities into copies of molecules, hence paving the path to follow “gene expression by numbers” (Rosenfeld et al., 2005).

In this article, we analyze gene expression following non-viral gene delivery, with focus on the variance of expression levels. The expression of genes exhibits all-or-nothing characteristics (Hume, 2000) and additional stochasticity exists in transcriptional regulation (McAdams and Arkin, 1999; Rao et al., 2002). Elowitz et al. (2002) have analyzed noise in bacterial gene expression and elucidated the distinction between extrinsic and intrinsic noise, that is, the contribution of fluctuations in cellular components and inherent stochasticity of the biochemical processes during gene expression. The extrinsic variance of gene expression within a clonal population of eukaryotic cells has been investigated in the light of stochastic theories (Blake et al., 2003; McAdams and Arkin, 1997; Raser and O’Shea, 2004; Volfson et al., 2006). It was only recently that attempts were made to generate models for transgene expression following non-viral gene delivery (Dinh et al., 2007; Varga et al., 2000, 2001; Zhou et al., 2007). Computational modeling might greatly enhance our understanding of gene transfer and aid in elucidating the nature of the underlying transport barriers. Many of the issues regarding cell entry and intracellular transport are shared with attempts to model viral infection (Douglas, 2008; Varga et al., 2005).

In this article, we used quantitative single cell time-lapse microscopy combined with mathematical modeling to analyze the variability in transgene expression (Fig. 1). From the synthetic delivery agents currently being evaluated for therapeutic use, we chose polyethyleneimine (PEI) (Boussif et al., 1995) and the commercial Lipofectamine 2000, as cationic polymer and lipid model systems, respectively. Both synthetic vectors are able to condense plasmid DNA into DNA-nano particles, denoted as cationic lipid (cationic polymer)–DNA complexes or just “complexes.” Distributions of the expression onset times and expression steady state levels were evaluated for both vectors.

Data are well described by a stochastic delivery model, which is based on the assumption that in a decisive step, only a small number of complexes enter the nucleus through a stochastic process. Out of these complexes, only a fraction of the plasmid load is expressed (Fig. 1). The theoretical model is further corroborated by a cotransfection analysis, that is, the case of the simultaneous transfection using two distinguishable plasmids encoding for CFP and YFP. It is shown that this model consistently describes the fraction of transfected cells and the observed expression level distribution. As a consequence the effective size of a stochastically delivered unit of plasmids (complex) can be determined.

## Materials and Methods

### Cell Culture

A human bronchial epithelial cell line (BEAS-2B, ATCC) was grown in Earle’s MEM supplemented with 10% FBS at 37°C in a humidified atmosphere, 5% CO<sub>2</sub> level. Transfection was performed on both non-synchronized and synchronized cultures. A thymidine kinase double-block was performed to synchronize cells.

### Transfection

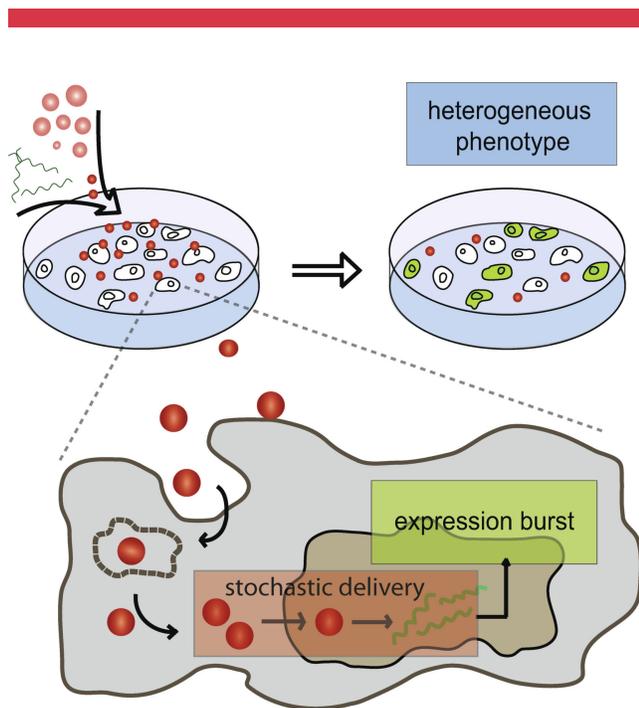
BEAS-2B cells were grown to 80% confluence from an initial seeding density of  $1 \times 10^5$  cells/well in six-well plates 24 h before transfection. Cells were washed and the medium is replaced with 1 mL OptiMEM/well immediately before transfection. Optimized transfection procedures were performed using either 2% (v/v) Lipofectamine<sup>TM</sup>2000/OptiMEM or PEI (N/P=8)/HBS; 1 μg of pEGFP-N1 or pd2EGFP-N1, is used for transfecting each batch of cells. The transfection medium was prepared either by adding the Lipofectamine or the PEI solution to the plasmid solution. After the transfection media were allowed to stand for 20 min the cells were incubated with 200 μL/well Lipofectamine or PEI transfection medium for 3 h at 37°C, 5% CO<sub>2</sub> level. After 3 h of incubation the medium was removed, and cells were washed with PBS. Cells were reincubated with Leibovitz’s L-15 Medium with 10% FBS prior to EGFP expression monitoring.

### Cotransfection

Cotransfection was performed with two kinds of preparations containing the same molar amount per plasmid. For one preparation, ECFP/Lipofectamine, EYFP/Lipofectamine, ECFP/PEI and EYFP/PEI were complexed separately. For the other, a mixture of ECFP and EYFP hetero-complexes were complexed with Lipofectamine or PEI. Transfection using either the hetero-complexes (pre-mixed) or a mixture of homo-complexes (post-mixed) was performed as previously described. Cells were reincubated in growth medium and CFP and YFP expression was monitored by fluorescence microscopy after 24 h.

### Data Acquisition and Quantitative Image Analysis

Images were taken at 10× magnification, with a constant exposure time of 1 s, at 10 min intervals for at least 30 h post-transfection. Fluorescence images are consolidated into single image sequence files. Negative control images were taken to assess lamp threshold values and autofluorescence, and were subtracted from corresponding image sequence files to eliminate autofluorescence effects. To capture cell



**Figure 1.** Experimental setup for single cell transfection experiments (**upper part**) and key elements of the theoretical model (**lower part**). EGFP-encoding plasmids and cationic agents form complexes, which are administered to eukaryotic cell cultures. Automated single cell microscopy yields statistics on phenotypic expression of EGFP. For the delivery of plasmids to the nucleus, stochastic effects are important, while the following expression of fluorescent proteins can be described in a deterministic fashion.

fluorescence over the entire sequence, regions of interest (ROIs) were manually defined around each cell (Fig. 2). Changes in total gray measurements in individual ROIs were determined for each time point.

## Results

### Time Lapse Microscopy and Single Cell EGFP Expression

A cell line of lung epithelium cells was transfected with a plasmid encoding for the green fluorescent protein (EGFP). Transfection protocols for PEI- and Lipofectamine-mediated delivery followed standard procedures and are described in detail in the Supplementary Online Information. We denote the time of the gene vector administration to the cell culture as  $t = 0$  h. Transfection medium was removed and cell growth medium added at  $t = 3$  h. Single-cell EGFP expression was monitored by automatically taking sequences of fluorescence micrographs from 25 view fields at 10-min intervals. Figure 2a shows a representative sequence from a Lipofectamine transfection experiment, with the initial bright field image, as well as the EGFP fluorescence at  $t = 4, 8,$  and  $12$  h post-transfection. These images demonstrate heterogeneity in both the

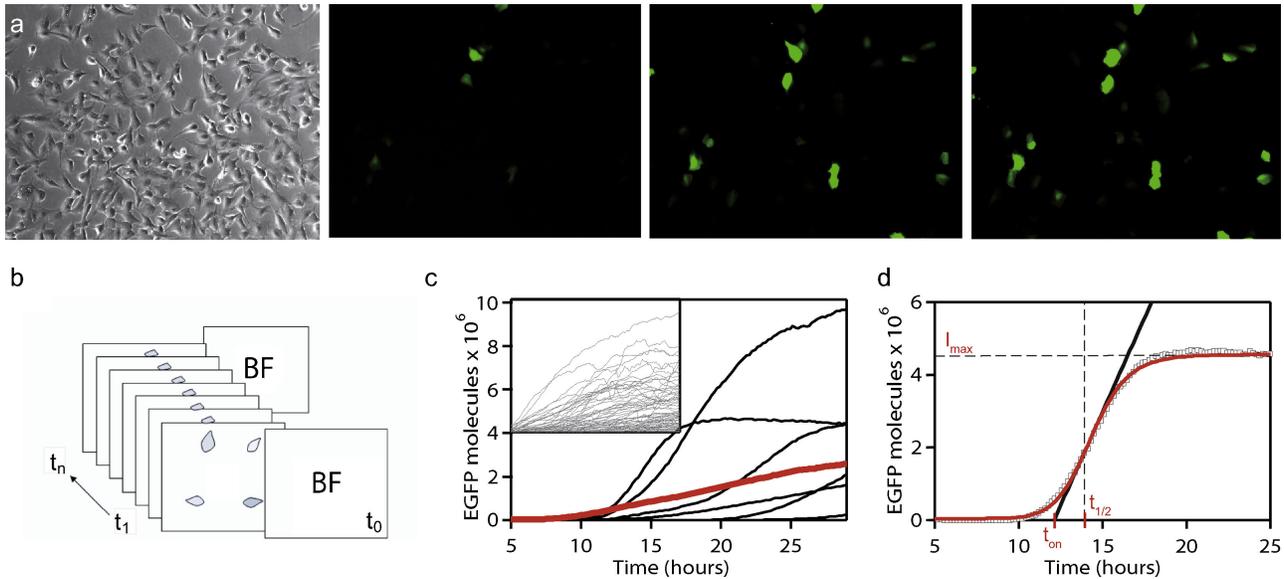
expression onset times and levels of exogenous gene expression. It is observed that the number of fluorescent cells increases with time; at the late stage ( $\sim 30$  h), the ratio of fluorescent to non-fluorescent cells is about 23% and 30% for PEI- and Lipofectamine-mediated transfection, respectively. A total of 500–1,500 cells were monitored in parallel within one time-lapse experiment. Individual time courses of the total fluorescence per cell were evaluated by image processing from data stacks as shown in Figure 2b and described in the Supplementary Online Information. Figure 2c shows a series of representative time traces from one transfection experiment, illustrating the significant variance in both the expression onset time and EGFP expression level. The typical sigmoidal shape of the time courses is well described by the phenomenological function

$$I(t) = \frac{I_{\max}}{2} \left[ 1 + \tanh\left(\frac{t - t_{1/2}}{t_{\text{rise}}}\right) \right] \quad (1)$$

which allows the determination of the maximal fluorescence plateau value ( $I_{\max}$ ) the time of half-maximum ( $t_{1/2}$ ) and the characteristic rise time,  $t_{\text{rise}}$ . The fluorescence intensities were converted into molecular units using EGFP standard beads for calibration (see Supplementary Online Information). In the remainder of the text we will give the fluorescence intensity  $I_{\max}$  in units of EGFP numbers  $G$ . Equation (1) proved to be robust for automated data analysis, facilitating the accumulation of statistics for a large number of individual cells. In order to determine the time points of expression onset,  $t_{\text{on}}$ , we use  $t_{1/2} - t_{\text{rise}}$  as an approximation due to the lack of a well-defined point of onset as shown in Figure 2d. Since the timing of gene expression is expected to be dependent on the cell cycle, we investigated the distribution of  $t_{\text{on}}$  for synchronized and non-synchronized cells. To this end, cells were arrested at the G1/S-Phase transition using a thymidine kinase double-block (Merrill, 1998).

### Distribution of Expression Onset Times

Figure 3 summarizes the measured distribution functions of  $t_{\text{on}}$  and  $I_{\max}$  for GFP expression after transfection of non-synchronized and synchronized cultures with PEI- and Lipofectamine-based complexes respectively. Expression onset times range from 5 and 25 h for both PEI- and Lipofectamine, indicating the existence of a time window during which plasmids are successfully transcribed in the nucleus. The distribution for Lipofectamine clearly peaks at an earlier time ( $\sim 8$  h) compared to PEI ( $\sim 16$  h). On cell cycle synchronization, the distribution of onset times sharpen and become peaked at about  $t = 15$  h for both PEI and Lipofectamine. Bright field images reveal that most synchronized cells divide 12 h after transfection. This is consistent with the fact that transfection was carried out in mid-S-phase, 3 h after release from the thymidine kinase



**Figure 2.** Acquisition of single cell time series. **a:** Microscopy view fields from a Lipofectamine transfection experiment. The first frame is a bright field (BF) control image. Fluorescence image sequences are taken automatically at 10-min intervals for at least 30 h. **b:** Definition of regions of interest (ROIs), total gray value measurement and conversion to the number of EGFP molecules. **c:** Representative time-courses of EGFP expression in individual cells following PEI-transfection. The population average (red) is plotted to demonstrate its linear increase in contrast to the sigmoidal shape of the individual traces. **d:** Characteristic parameters of expression are obtained by fitting the heuristic function 1 (red) to the recorded fluorescence time course (black). The time of expression onset,  $t_{on}$ , is calculated from the time of half-maximal expression  $t_{1/2}$  and the slope at that point.

double-block. This shows that plasmid activation occurs about 3 h after the M-phase. Furthermore, since cell cycle synchronization suppresses expression at earlier times, there is evidence that the delivery process depends on the cell cycle dependent breakdown of the nuclear membrane. This is consistent with previous studies claiming that mitosis enhances transgene nuclear translocation in cationic lipid gene delivery (Brunner et al., 2000; Mortimer et al., 1999; Tseng et al., 1999). We also find that synchronization leads to a twofold higher steady state expression for PEI-mediated (from  $2.6 \times 10^6$  to  $4.3 \times 10^6$  average EGFP molecules per cell) and Lipofectamine-mediated (from  $2.9 \times 10^6$  to  $5.1 \times 10^6$  average EGFP molecules per cell) transfection, consistent with earlier observations on ensemble averaged data (Brunner et al., 2000).

### Modeling Steady State Gene Expression

In order to analyze the distribution of expression steady states, we introduce a mathematical model that describes EGFP expression after transfer and nuclear translocation of complexes containing exogenous EGFP plasmids. Stochasticity due to nuclear translocation of the plasmid complexes and the intra-nuclear activation will give rise to a probability distribution  $P(X)$  for  $X$  successfully expressed plasmids (see Fig. 4a). In the first stage of analysis, we describe the expression of EGFP from a single activated plasmid in a linear deterministic model and neglect any

cell-cell variability. Based on the biochemical reactions describing transcription, translation and maturation as shown in Figure 4b we denote the ensuing rate equations. Here,  $R$  denotes the number of RNA molecules,  $U$  the number of unfolded polypeptide chains, and  $G$  the number of folded EGFP proteins:

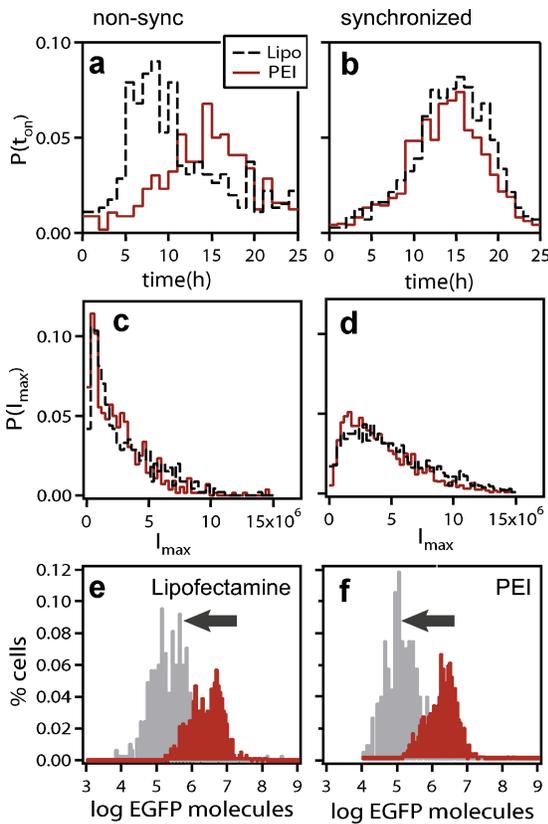
$$\dot{R} = s_A X - \delta_R R \quad (2)$$

$$\dot{U} = s_p R - (k_M + \delta_U) U \quad (3)$$

$$\dot{G} = k_M U - \delta_G G \quad (4)$$

$s_A$ ,  $s_p$ , and  $k_M$  denote the rate constants for transcription, translation and EGFP maturation,  $\delta_R$ ,  $\delta_U$ , and  $\delta_G$  denote the degradation constants of each product, respectively. The degradation rates of folded ( $\delta_G$ ) and unfolded protein ( $\delta_U$ ) are assumed to be equal, since the same proteases are involved (Leveau and Lindow, 2001). A plasmid degradation term was omitted, since its occurrence is predicted to be negligible within the time frame considered (Subramanian and Sreenc, 1996). Literature values for the individual kinetic rates are summarized in Table I. Equations (2)–(4) can be solved analytically. For the steady state value, a linear relation

$$I_{max} = G(t \rightarrow \infty) = \frac{k_M s_p s_A}{\delta_G \delta_R (k_M + \delta_G)} X \quad (5)$$



**Figure 3.** EGFP expression statistics for PEI- and Lipofectamine-mediated transfection. Distributions of expression onset times  $t_{on}$  (a and b) and maximal expression values  $I_{max}$  (c and d), for PEI-mediated (red) and Lipofectamine-mediated (dashed black) transfection depict strong variability within the cell cultures. The total number of expressing cells was 23% out of 560 for PEI and 30% out of 502 in the case of Lipofectamine. b and d: Thymidine kinase-synchronized cultures with 40% out of 1981 and 30% out of 1797 cells expressing EGFP for PEI and Lipofectamine, respectively. For synchronized cells, expression onset time distributions coincide for Lipofectamine and PEI, indicating that transfection is more likely to happen in specific phases of the cell cycle. Distributions for  $I_{max}$  (given in units of EGFP molecules) cannot be explained by post-transfectional sources of fluctuations alone. e and f: Effect of the altered expression rates on the distribution of maximal expression levels  $I_{max}$ . Distributions for d2EGFP (gray) and EGFP (red) transfected with Lipofectamine (e) or PEI (f) are shown. d2EGFP, which has a higher degradation rate, exhibits a systematic shift of the  $I_{max}$  distribution compared to EGFP, independent of the vector used. Besides this shift, a change in the number of proteins per active plasmid,  $k_{exp}$ , preserves the shape of the distribution. This suggests that the shape is determined during plasmid delivery prior to expression.

between the expression level  $I_{max} = [GFP]$  and the number of expressed plasmids  $X = [\text{plasmids}]$  is obtained:

$$[GFP] = k_{exp}[\text{plasmids}] \quad (6)$$

Here  $k_{exp}$  denotes an effective expression factor, corresponding to the number of proteins expressed per transcribed plasmid in the steady state. With the values given in Table I, we find  $k_{exp} \approx 4 \times 10^6$  molecules/plasmid, which compared to the experimental number of molecules ( $1-15 \times 10^6$ ), results in the remarkable finding that the number of plasmids  $X$  is of order 1. This implies that most of the

variance in expression level originates from stochastic variations in the small number of plasmids, such that the distribution of GFP expression is determined by the distribution of successfully delivered plasmids,  $P(G) \sim P(X)$ .

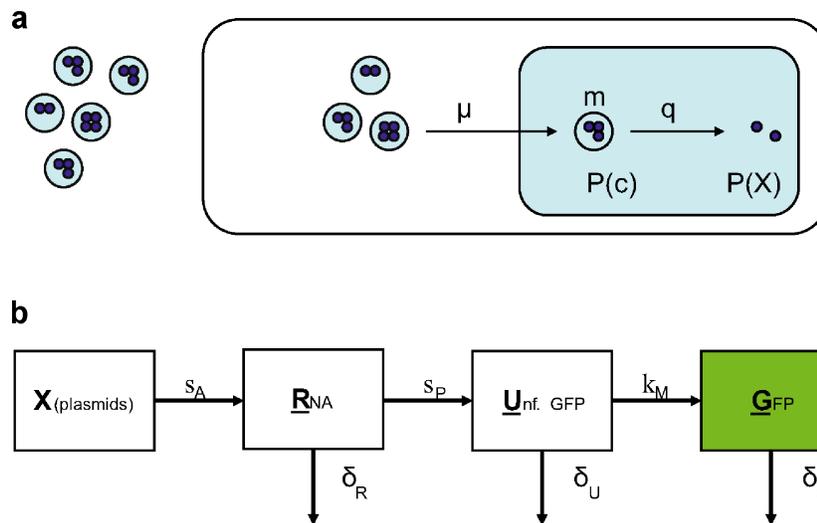
To further substantiate this conclusion, we designed an experiment where the expression factor  $k_{exp}$  is deliberately modified through the use of destabilized EGFP. It has a 14-fold higher degradation rate ( $\delta_{desG}$ ) due to an additional amino acid sequence (PEST), which makes it more susceptible to proteolysis (Kain, 1999). Figure 3e and f display the shift in the steady state distribution of  $I_{max}$ , shown in a logarithmic scale. As predicted above, the shape of the distribution function is almost unchanged for both PEI- and Lipofectamine-mediated transfection. In addition, the peak positions shifted by a factor 12.5, which is close to the value 14.3 predicted from Equation (5).

### Modeling Transfection Noise

Unlike in chromosomal DNA, which contains a fixed number of genes, the transfection experiments discussed here result in the delivery of a variable number of genes per vector. We model gene delivery as a two-step stochastic process as shown in Figure 4a. As we will argue in the following, a two-step model is the simplest model that is in accordance with the experimental data. The model consists of (i) the nuclear translocation with probability  $\mu$  of complexes containing an average of  $m$  plasmids and (ii) intra-nuclear activation of plasmids, with probability  $q$ . Whereas the probability  $q$  subsumes all phenomena promoting or interfering with transcription such as DNA methylation or complexation. We assume that the first process, the delivery of complexes to the nucleus, is rare and statistically independent, yielding a Poisson distribution for the number of delivered complexes  $C$ :

$$P(C) = \frac{\mu^C}{C!} e^{-\mu} \quad (7)$$

characterized by its mean value  $\mu$ . Secondly, the independent activation of a plasmid in the nucleus is described by a Bernoulli process with success probability  $q$ . The concatenation of both processes results in an expression for  $P(X)$  which retains the characteristics of a Poissonian. Mathematical details of its derivation can be found in the Supplementary Data. Figure 5 shows the calculated distribution of activated plasmids,  $P(X)$  (red bars), to the measured experimental protein distribution,  $P_{exp}(G)$  (green bars). In addition a theoretical protein distribution is shown as black lines.  $P_{theo}(G)$  is obtained from  $P(X)$  by additionally accounting for noise in gene expression, where we have used a relative magnitude of 0.3 for post-transfectional noise from the literature (see Supplementary Data). Note that the  $x$ -axis of the distributions  $P(G)$  is rescaled by the factor  $k_{exp}$  according to Equation (6). The agreement between



**Figure 4.** Theoretical model for transfection and gene expression. **a:** Our model of plasmid delivery consists of several stochastic components. The number of complexes  $C$  delivered per cell is Poisson-distributed, with mean  $\mu$ . Each complex carries a random number of plasmids, described by a Poisson distribution with mean  $m$ . Finally, each plasmid has an activation probability  $q$ , resulting in a Binomial distribution of active plasmids  $X$  out of the total number of delivered plasmids. With this approach, the overall distribution,  $P(X)$ , of actively expressing plasmids can be derived. **b:** Deterministic model of EGFP expression including transcription ( $s_A$ ), translation ( $s_P$ ) and protein maturation ( $k_M$ ). mRNA ( $R$ ), unfolded proteins ( $U$ ) and GFP ( $G$ ) are degraded with rates  $\delta_R$ ,  $\delta_U$  and  $\delta_G$ , respectively. Solving the corresponding rate equations, the steady state distribution of fluorescent proteins,  $P(G)$ , can be related to that of active plasmids,  $P(X)$ .

experiment and model is remarkable considering that there is only one free parameter in the fit. This is due to the fact that two additional experimental constraints have to be met. These are the measured fraction of transfected cells, TR, defined as the percentage of cells expressing one or more plasmids, and the average number of GFP molecules per cell,  $\langle G \rangle$ , determined by calibration. The parameters  $\mu$  and  $m$  are fixed by these constraints. The remaining unknown is the expression factor,  $k_{exp}$ , which is determined by the fit shown in Figure 5. We obtain  $k_{exp} \approx 1 \times 10^6$ ,  $m_{eff} \approx 3$ , and  $\mu \approx 0.3-0.5$ .

The red curve in Figure 5 represents the distribution of successfully expressed plasmids. This distribution, which is directly related to the number of expressed GFP protein though Equation (6), has a well defined mean value given by

$$\langle [\text{plasmids}] \rangle = \mu m q \quad (8)$$

The transfection ratio, TR, is related to this mean plasmid number. It depends on the average number of complexes

delivered  $\mu$  and the effective probability  $\tilde{q}$  that from any given complex at least one plasmid is transcribed (we present a complete derivation of these quantities in the Supplementary Data).

$$\text{TR}(\mu, m, q) = 1 - \exp\{-\mu \tilde{q}\} \quad (9)$$

For the data shown in Figure 6, TR is of order 20%.

### Cotransfection and Correlated Delivery

One important ingredient of our model is the delivery of DNA in units or complexes and the subsequent correlated coexpression of multiple plasmid copies. This assumption is closely related to the question of whether DNA-complexes fully dissociate before nuclear entry or complexes enter the nucleus as a whole. To elucidate this issue, we studied cotransfection of two distinguishable plasmids (CFP and YFP) and analyzed the outcome of transfection using pre-mixed and post-mixed complexes. Pre-mixed complexes

**Table I.** Literature values for the kinetic rates of the linear gene expression model

| Rate   | Best estimate in $\text{h}^{-1}$ | Literature values in $\text{h}^{-1}$ | References                                       |
|--|----------------------------------|--------------------------------------|--|
| $s_A$ (transcription)                              | 180                              | 145–240                              | Hume (2000)                                      |
| $s_P$ (translation)                                | 100                              | 60–180                               | Molecular Biology of the Cell 3rd Edition (1995) |
| $k_M$ (fluorophore maturation)                     | 1.0                              | 0.96–1.28                            | Sniegowski et al. (2005)                         |
| $\delta_M$ (mRNA degradation)                      | 0.10                             | <0.14                                | Sacchetti et al. (2001)                          |
| $\delta_P$ (protein degradation)                   | 0.035                            | 0.03–0.04                            | Sacchetti et al. (2001)                          |
| $\delta_{desG}$ (destabilized protein degradation) | 0.5                              | 0.5                                  | BD Biosciences Clontech                          |

contain both CFP- and YFP-plasmids in a single complex, whereas post-mixed complexes contain either CFP- or YFP-plasmids (for clarity see Fig. 6a and b). The steady-state CFP/YFP expression was analyzed at 24 h post-transfection for approximately 15,000 cells. We define the cotransfection ratio,  $r$ , as the number of cells expressing both CFP and YFP divided by the number of cells expressing either CFP or YFP. We find that the cotransfection ratio increases from 12.9% for post-mixed complexes to 21.9% for premixed complexes. The significant difference could not be explained, if complexes were completely dissolved in the cytosol and delivery of plasmids was independent from the complexes. The two-step delivery model, however, naturally explains the discrepancy between pre-mixed and post-mixed complexes. Based on our model, an analytical expression for the cotransfection ratio can be derived (see Supplementary Data) which predicts correctly the measured cotransfection

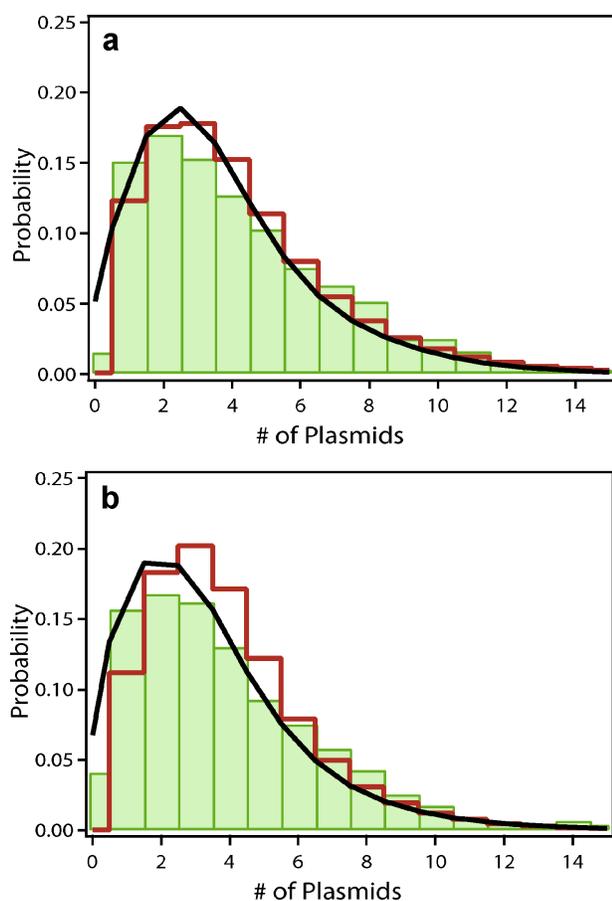
ratios, if the same parameters are used as determined from the EGFP distribution function.

## Discussion

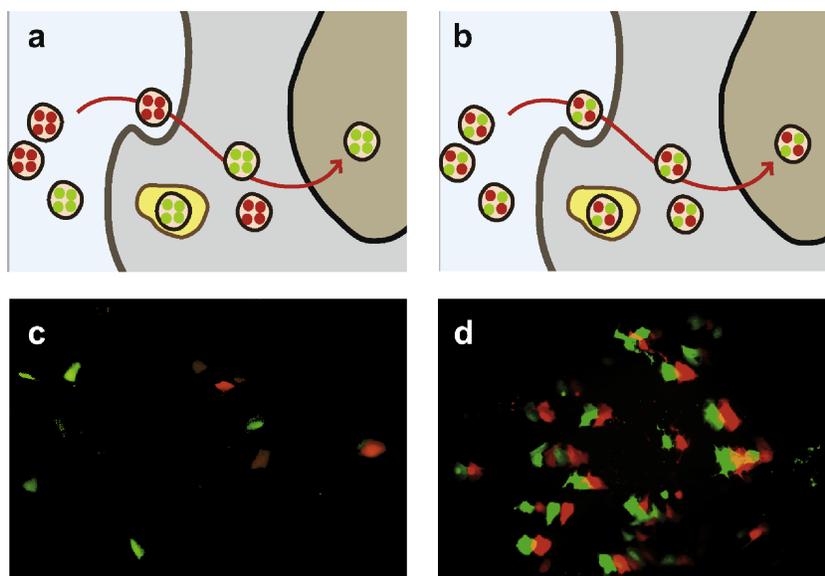
We have measured the distribution of expression onset times and steady-state expression levels derived from single cell fluorescence time courses. Distributions of onset times of PEI and Lipofectamine collapse on a single curve for synchronized cell cultures, suggesting a universal cell cycle-dependent gene delivery mechanism. Synchronized cells exhibit a broad Poissonian distribution in expression levels and cotransfection experiments reveal correlations in the delivery probability for plasmids contained in one complex.

Invoking Occam's razor, we analyzed the findings in terms of an idealized minimalist model of gene transfection, which describes gene delivery as a two-step stochastic process. Yet our model proves to have considerable predictive power by relating measurable quantities such as the overall transfection efficiency, the cotransfection probability and the shape of the gene expression distribution with each other. Thus, the model allows the derivation of the expression factor, the number of activated plasmids per complex and the average number of delivered complexes from the measured single cell transfection statistics. The model also elucidates the origin of expression variance, separating the noise due to small number fluctuations of complexes, which is inherent to the delivery process and extrinsic sources of noise due to cell-cell variability.

In our gene expression model, we refer to complexes as units of coherently delivered plasmids. Those indirectly inferred complexes are consistent with but not necessarily identical to the complexes described in many physico-chemical studies of PEI and lipofectamine mediated transfection. Cationic-lipid complexes are known to form multi-lamellar aggregates that contain a large number of plasmids (Lasic et al., 1997; Rädler et al., 1997; Zabner et al., 1995). Following endocytotic uptake and release, the complexes slowly dissociate in a stepwise, unwrapping mechanism (Kamiya et al., 2002; Lin et al., 2003). PEI complexes are torroids or rods with a typical hydrodynamic radius of 100 nm (Boussif et al., 1995; DeRouchey et al., 2005), which have been seen to be actively transported inside cells (de Bruin et al., 2007) and to accumulate in the periphery of the nucleus (Suh et al., 2003). Both scenarios describe a situation where numerous small complexes have equal chances of entering the nucleus during the course of mitosis, which is consistent with our model assumptions. Microscopy studies have argued favorably for complexes being at least not fully dissolved at the final delivery stage (Lin et al., 2003; Tseng et al., 1999). However, single nuclear entry events have not been documented explicitly. The probability of transgene expression in the nucleus again depends on the nature of the transfection agent. Pollard et al. (1998) reported that cationic lipids, but not PEI prevent gene expression when complexes are directly injected in the



**Figure 5.** Comparison of single-cell data with the theoretical model. The theoretical EGFP distribution (black) is intimately connected with the underlying distribution of expressing plasmids (red). To facilitate comparison, the protein distribution has been scaled down by the average number of proteins per active plasmid in steady state,  $k_{exp}$ . **a** and **b**: For synchronized cultures the measured protein distribution (green) is fitted very well by our theoretical model (black). The fit for PEI transfection (**a**) yields an average number of delivered complexes,  $\mu = 0.53$ , and an average number of activated plasmids per complex,  $m_{eff} = 3.2$ . In the case of Lipofectamine (**b**), we find  $\mu = 0.37$  and  $m_{eff} = 3.2$ .



**Figure 6.** Correlated delivery in CFP/YFP cotransfection with post-mixed (uni-colored) complexes (left) and pre-mixed (dual-colored) complexes (right). **a** and **b**: Post-mixed (uni-colored) and pre-mixed (dual-colored) complexes carry different plasmid content, but take the same pathway to the nucleus. **c** and **d**: Superposition of CFP and YFP fluorescence after transfection reveals a qualitatively different expression pattern for the two distinct experimental protocols. Cyan fluorescence is slightly displaced to permit identification of cotransfected cells. All micrographs are artificially colored.

nucleus. Such findings can only be consolidated with our model, if the delivered complexes transform during the course of the delivery, rather than being the same physical complexes as originally prepared under *in vitro* conditions. Within the context of our model we restrict ourselves to a narrowed meaning of “complexes” as units of plasmids that are co-delivered. In this framework, we determine the average number of successfully delivered complexes and the effective number of activated plasmids per complex from the analysis of single cell statistics. It will be interesting to corroborate the physical fate and the expression outcome of single complexes by high resolution studies in single cells (de Bruin et al., 2007).

The method to use transfection assays based on automated high-throughput microscopy combined with image processing might evolve into a routine tool for the assessment of transfection efficiency. In contrast to ensemble averaged fluorescence or luminescence data, single cell assays yield precise distribution functions and single cell expression dynamics, which allow a more detailed comparison to theoretical models. As shown here, the analysis of steady state expression levels provides access to the probability of successful plasmid delivery ( $P(X)$ ) and yields an absolute number for the expression factor ( $k_{\text{exp}}$ ). In forthcoming work we will discuss in more detail the distribution of expression onset times and the expression dynamics. We expect that our particular mathematical model can be adapted to a wider class of transfection agents and different types of cells. Their distinct transfection ratios, rate constants and numbers of effective complexes will become even more meaningful in the context of

comparative theoretical modeling. A combined experimental and modeling approach will hence help to identify rate-limiting barriers to gene transfer and will result in improved data comparability, making it a versatile tool in the continuous evaluation and improvement of existing synthetic vectors.

We thank Josef Rosenecker, Carsten Rudolph and Ernst Wagner for fruitful discussions and Susanne Kempfer for continuous support. The work was supported by the Deutsche Forschungsgemeinschaft through grant SFB486-B10 and SFB TR12, through the Excellence Cluster “Nanosystems Initiative Munich (NIM),” and through the LMU innovative program “Analysis and Modelling of Complex Systems.” S.Y., J.-T.K. and M.P.D. gratefully acknowledge scholarships from Microsoft Research, the IDK NanoBioTechnology and the DAAD, respectively.

## References

- Blake WJ, Mads KA, Cantor CR, Collins JJ. 2003. Noise in eukaryotic gene expression. *Nature* 422(6932):633–637.
- Boussif O, Lezoualch F, Zanta MA, Mergny MD, Scherman D, Demeneix B, Behr JP. 1995. A versatile vector for gene and oligonucleotide transfer into cells in culture and *in vivo* polyethylenimine. *Proc Natl Acad Sci USA* 92(16):7297–7301.
- Brunner S, Sauer T, Carotta S, Cotten M, Saltik M, Wagner E. 2000. Cell cycle dependence of gene transfer by lipoplex polyplex and recombinant adenovirus. *Gene Ther* 7(5):401–407.
- de Bruin K, Ruthardt N, von Gersdorff K, Bausinger R, Wagner E, Ogris M, Braeuchle C. 2007. Cellular dynamics of EGF receptor-targeted synthetic viruses. *Mol Ther* 15(7):1297–1305.
- DeRouchev J, Netz RR, Radler JO. 2005. Structural investigations of DNA-polycation complexes. *Eur Phys J E Soft Matter* 16(1):17–28.

- Dinh A, 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(3):831–846.
- Douglas KL. 2008. Toward development of artificial viruses for gene therapy: A comparative evaluation of viral and non-viral transfection. *Biotechnol Prog* 24(4):871–883.
- Elowitz MB, Levine AJ, Siggia ED, Swain PS. 2002. Stochastic gene expression in a single cell. *Science* 297(5584):1183–1186.
- Ferber D. 2001. Gene therapy: Safer and virus-free? *Science* 294(5547):1638–1642.
- Hume DA. 2000. Probability in transcriptional regulation and its implications for leukocyte differentiation and inducible gene expression. *Blood* 96(7):2323–2328.
- Kain S. 1999. Green fluorescent protein (GFP): Applications in cell-based assays for drug discovery. *Drug Discov Today* 4(7):304–312.
- Kamiya H, Fujimura Y, Matsuoka I, Harashima H. 2002. Visualization of intracellular trafficking of exogenous DNA delivered by cationic liposomes. *Biochem Biophys Res Commun* 298(4):591–597.
- Kirchies R, Wightman L, Wagner E. 2001. Design and gene delivery activity of modified polyethylenimines. *Adv Drug Deliv Rev* 53(3):341–358.
- Lasic D, Strey H, Stuart M, Podgornik R, Frederik P. 1997. The structure of DNA-liposome complexes. *J Am Chem Soc* 119(4):832–833.
- Lechardeur D, Verkman AS, Lukacs GL. 2005. Intracellular routing of plasmid DNA during non-viral gene transfer. *Adv Drug Deliv Rev* 57(5):755–767.
- Leveau JHJ, Lindow SE. 2001. Predictive and interpretive simulation of green fluorescent protein expression in reporter bacteria. *J Bacteriol* 183(23):6752–6762.
- Lin A, Slack N, Ahmad A, George C, Samuel C, Safinya C. 2003. Three-dimensional imaging of lipid gene-carriers: Membrane charge density controls universal transfection behavior in lamellar cationic liposome-DNA complexes. *Biophys J* 84(5):3307–3316.
- Longo D, Hasty J. 2006. Dynamics of single-cell gene expression. *Mol Syst Biol* 2(64):1–10.
- McAdams HH, Arkin A. 1997. Stochastic mechanisms in gene expression. *Proc Natl Acad Sci USA* 94(3):814–819.
- McAdams HH, Arkin A. 1999. It's a noisy business! Genetic regulation at the nanomolar scale. *Trends Genet* 15(2):65–69.
- Merrill G. 1998. Cell Synchronization. *Methods Cell Biol* 57:229–249.
- Mortimer I, Tam P, MacLachlan I, Graham RW, Saravolac EG, Joshi PB. 1999. Cationic lipid-mediated transfection of cells in culture requires mitotic activity. *Gene Ther* 6(3):403–411.
- Patil SD, Rhodes DG, Burgess DJ. 2005. DNA-based therapeutics and DNA delivery systems: A comprehensive review. *AAPS J* 7(1):E61–E77.
- Pollard H, Remy J, Loussouarn G, Demolombe S, Behr J, Escande D. 1998. Polyethylenimine but not cationic lipids promotes transgene delivery to the nucleus in mammalian cells. *J Biol Chem* 273(13):7507–7511.
- 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(5301):810–814.
- Rao CV, Wolf DM, Arkin AP. 2002. Control, exploitation and tolerance of intracellular noise. *Nature* 420(6912):231–237.
- Raser JM, O'Shea EK. 2004. Control of stochasticity in eukaryotic gene expression. *Science* 304(5678):1811–1814.
- Rosenfeld N, Young JW, Alon U, Swain PS, Elowitz MB. 2005. Gene regulation at the single-cell level. *Science* 307(5717):1962–1965.
- Roth CM, Sundaram S. 2004. Engineering synthetic vectors for improved DNA delivery: Insights from intracellular pathways. *Annu Rev Biomed Eng* 6:397–426.
- Sacchetti A, El Sewedy T, Nasr AF, Alberti S. 2001. Efficient GFP mutations profoundly affect mRNA transcription and translation rates. *FEBS Lett* 492(1-2):151–155.
- Safinya CR. 2001. Structures of lipid-DNA complexes: Supramolecular assembly and gene delivery. *Curr Opin Struct Biol* 11(4):440–448.
- 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. *J Biol Chem* 280(28):26248–26255.
- Subramanian S, Srien F. 1996. Quantitative analysis of transient gene expression in mammalian cells using the green fluorescent protein. *J Biotechnol* 49(1-3):137–151.
- Suh J, Wirtz D, Hanes J. 2003. Efficient active transport of gene nanocarriers to the cell nucleus. *Proc Natl Acad Sci USA* 100(7):3878–3882.
- Tseng W, Haselton F, Giorgio T. 1999. Mitosis enhances transgene expression of plasmid delivered by cationic liposomes. *Biochim Biophys Acta* 1445(1):53–64.
- Varga C, Wickham T, Lauffenburger D. 2000. Receptor-mediated targeting of gene delivery vectors: Insights from molecular mechanisms for improved vehicle design. *Biotechnol Bioeng* 70(6):593–605.
- Varga C, Hong K, Lauffenburger D. 2001. Quantitative analysis of synthetic gene delivery vector design properties. *Mol Ther* 4(5):438–446.
- Varga CM, Tedford NC, Thomas M, Klivanov AM, Griffith LG, Lauffenburger DA. 2005. Quantitative comparison of polyethylenimine formulations and adenoviral vectors in terms of intracellular gene delivery processes. *Gene Ther* 12(13):1023–1032.
- Volfson D, Marciniak J, Blake W, Ostroff N, Tsimring L, Hasty J. 2006. Origins of extrinsic variability in eukaryotic gene expression. *Nature* 439(7078):861–864.
- Zabner J, Fasbender AJ, Moninger T, Poellinger KA, Welsh MJ. 1995. Cellular and molecular barriers to gene-transfer by a cationic lipid. *J Biol Chem* 270(32):18997–19007.
- Zhou J, Yockman JW, Kim SW, Kern SE. 2007. Intracellular kinetics of non-viral gene delivery using polyethylenimine carriers. *Pharm Res* 24(6):1079–1087.



# Appendix B

## Original publication P2

Maria Pamela David, Johnrob Bantang, Eduardo Mendoza

**A Projective Brane Calculus with Activate, Bud and Mate as Primitive Actions**  
in C. Priami et al. (Eds.): *Trans. on Comput. Syst. Biol. XI, LNBI*, **5750**: 164 - 186,  
2009.



# A Projective Brane Calculus with Activate, Bud and Mate as Primitive Actions

Maria Pamela C. David<sup>1,\*</sup>, Johnrob Y. Bantang<sup>1,2,3,\*</sup>,  
and Eduardo R. Mendoza<sup>1,4</sup>

<sup>1</sup> Faculty of Physics and Center for Nanoscience, Ludwig-Maximilians-Universität München, Geschwister-Scholl-Platz 1, D-80539 München, Germany

<sup>2</sup> Max-Planck-Institut für Dynamik komplexer technischer Systeme, Sandtorstraße 1, D-39106 Magdeburg, Germany

<sup>3</sup> National Institute of Physics, College of Science, University of the Philippines, Diliman, Quezon City 1101 Philippines

<sup>4</sup> Department of Computer Science, University of the Philippines, Diliman, Quezon City 1101 Philippines

**Abstract.** We modify and extend Cardelli’s Brane Calculus and Danos and Pradalier’s Projective Brane Calculus (PBC) to improve consistency with biological characteristics of membrane reactions. We propose a Projective Activate-Bud-Mate (PABM) calculus as an alternative to the Phago-Exo-Pino (PEP) basic calculus of L. Cardelli. PABM uses a generalized formalism for Action activation with receptor-ligand type channel construction that incorporates multiple association and affinity similar to Priami’s beta binders. Calculus elements are finite. Volumes are associated with systems for more realistic compartment-based reaction probabilities. PABM also uses *Brane domains* that partition membranes into controllable, independent groupings of projective actions. Domains eliminate the need for parameters in Phago and Bud and allow lateral and cross-membrane interactions. We show that PABM can emulate bitonal membrane reactions. PABM also realizes the idea of L. Cardelli (Cardeli, 2004) on modeling molecules as systems.

## 1 Introduction

Cellular organization plays a key role in biological systems through the physical regulation of reactions. Enzymes, for instance, are typically sequestered in membrane-bound systems to which access is only made possible through cascades of equally regulated and timed signals. Most current formalisms for modeling, however, do not possess an explicit functionality for modeling compartmentalization. In deterministic models, compartmentalization is modeled with the use of additional variables that differentiate a species  $S$  that is within some compartment  $X$  from  $S$  that is within another compartment  $Y$ . While this has been used with some success,  $S$  in  $X$  is actually not different from  $S$  in  $Y$ , unless it has already reacted with other species in either compartment. It has only been in

---

\* The first two authors contributed equally to this work.

recent years that several calculi were developed so that membrane compartmentalization: (a) becomes an inherent part of computations and (b) is emphasized in simulating reactions[1, 2, 3, 4].

Brane calculus is a formalism that can be used to describe systems as mem-brane-bound compartments that may contain other systems[3]. These compartments can merge, split or be hierarchically reorganized through uptake (phagocytosis) or extrusion (exocytosis) mechanisms, based on the capabilities, known as *actions*, of the membranes that enclose them[3, 5]. An important aspect of these actions — directly adapted from pi calculus — is that they are triggered via highly specific channel-based communication. Nevertheless, the mapping between channels is not necessarily one-to-one, with some channels having more than one communication partner. Although the original concept of channels in pi calculus was for mobile telecommunication systems, it is compatible with the representation of biological interactions, from enzyme-substrate systems that interact to form a chemically distinct product to receptor-mediated intermembrane communication that leads to membrane reorganization.

Another formalism that includes compartments is Priami’s beta binders[4]. Here, much emphasis is given to the promiscuousity of the channels (“beta binders”) through which the compartments interact, as well as its effects on the dynamic evolution of the compartment contents, interactions, and interfaces. As in Brane calculus, compartments can merge and split as a result of binder-based communication. While inherent in Brane Calculus, beta binders needed an extension to include hierarchical construction of compartments. Recent extensions, however, only permit intuitive representation of static hierarchical structures, but still forbid the explicit nesting of compartments[6]. The main advantage of beta binders over Brane calculus is its natural representation of affinity to channel pairings, a concept that is adapted in the proposed extension in this paper.

The uniqueness of Cardelli’s Brane calculus lies in the representation of all computations *on* membranes. This is important, particularly since it is actually the dynamic property of membranes that determines its capability to interact with other membranes and its general environment *in vivo*. Consequently, this property also determines how a membrane-bound system would evolve. Structure hierarchy can likewise be easily represented in Brane calculus, where nested systems are effectively organized in tree structures[3, 5].

Brane calculus has been previously extended by Danos and Pradalier to incorporate the idea that the inner and outer surfaces of a membrane are not identical[5]. *In vivo*, it could be frequently observed that the membrane protein domains exposed to the extracellular matrix are different from the domains exposed to the cytosol. It is even possible for membrane proteins to possess either an extracellular domain or a cytosolic domain. As a result, the definition of the inner and outer membranes are different. Additional physical restrictions are introduced on which reactions could take place, in particular only directed actions on membrane surfaces that could “see” each other are allowed to interact. This extension using directed actions is known as *projective Brane calculus* (PBC).

Nevertheless, there are a number of aspects in both calculi that involve concepts not observed in biological system. The purpose of this paper is to introduce further modifications and extensions combining the strength of both Brane calculi, with the aim of making it even more consistent with the biological characteristics of membrane reactions. Specifically, we introduce the following changes that result to the proposed extension, the Projective Activate-Bud-Mate calculus (PABM):

1. Use of the minimal set  $\mathcal{S}_{\min} \equiv \{\text{bud}, \text{mate}, !, 0\}$  instead of the set  $\mathcal{S} \equiv \{\text{phago}, \text{exo}, \text{pino}, \text{mate}, \text{bud}, \text{drip}\}$  for the possible actions  $a \in \mathcal{S}_{\min}$  (see §2). All other actions in  $\mathcal{S}$  are realized using only the actions in  $\mathcal{S}_{\min}$  together with directed Actions of PBC.
2. Abstraction of specific send-receive channel pairing into less specific channel name equality, eliminating the distinction between input and output channels. Together with the previous revision, it allows generalized representations in the form  $ax$  for Actions and Coactions, where  $x$  is a named channel.
3. Introduction of *Brane domains*, which are autonomous groups of directed Actions within a Brane. The use of Brane domains would also allow inter-domain interactions within and across the same membrane.
4. Removal of the parameters for Bud and Phago, allowing the dynamic nature of membranes to be reflected in the calculus.
5. Inclusion of volume information as a system attribute to reflect its effects on the probability at which collisions will occur inside a compartment.
6. Association of rates to channels emulating an affinity feature similar to beta binders.
7. Treatment of Brane constituents and contents as finite quantities.
8. Elaboration of molecules as systems, a concept previously introduced by L. Cardelli[3].

These modifications are also geared towards the development of a machine for Brane calculus that can handle large-scale biological models.

## 2 Modified Notations

Table 1 summarizes the proposed notation and conceptual changes to the current Brane calculus, provided as a quick reference to the detailed explanations for these changes in the succeeding sections.

### 2.1 Actions

**Notations and terms.** In the documentation for the design of a machine for Brane calculus[7], stochastic pi calculus notations for input and output channels are used to distinguish between actions and coactions. At this point, it is important to make a distinction between an action (small ‘a’),  $a$  and an Action (capital ‘A’),  $\sigma$ . An action is an element of the set,  $a \in \mathcal{S}$  currently defined as:

$$\mathcal{S} \equiv \{\text{phago}, \text{pino}, \text{exo}, \text{mate}, \text{bud}, \text{drip}\}; \quad (1)$$

**Table 1.** Comparison of the currently-established Brane calculus (Cardelli’s Brane calculus and PBC) with the proposed calculus. Note that  $a, a_i \in \mathcal{S}$   $\sigma, \sigma_i, \tau \in \mathcal{A}$ ,  $i = 1, 2$ , with  $\bar{a}$  as the coaction of  $a$ . Conventions for parallel composition from PBC [5] are used.

| <i>Definition</i>                         | <i>Brane/Pi Calculus</i>  | <i>PABM</i>  |
|---|---|--|
| Channel ( $x \in \mathcal{C}$ )           | $!x \longleftrightarrow ?x$   | $x \longleftrightarrow x$  |
| Set of actions ( $a \in \mathcal{S}$ )    | $\mathcal{S} \equiv \{\text{phago, pino, exo, mate, bud, drip, } \dots\}$ | $\mathcal{S}_{\min} \equiv \{0, \text{bud, mate, } !\}$  |
| Action ( $\sigma, \tau \in \mathcal{A}$ ) | $a!x \longleftrightarrow \bar{a}?x$                                       | $!x \longleftrightarrow ax$ ( $a \neq !$ )   |
| Brane domain                              | undefined   | $\rho \equiv \langle \sigma_1; \sigma_2 \rangle$   |
| Brane                                     | $\langle \sigma_1; \sigma_2 \rangle$                                      | $[\rho]$   |
| Directionality                            | $\sigma_1$ is outside, $\sigma_2$ is inside                               | $\sigma_1$ is outside, $\sigma_2$ is inside  |
| System                                    | $\langle \sigma_1; \sigma_2 \rangle \langle P \rangle$                    | $[\rho] \langle P \rangle$   |
| Parameter (bud and phago)                 | $\sigma(\tau)$  | $\tau.\sigma$ and $\rho$ (see text for details)  |
| Choice                                    | $\sigma_1 + \sigma_2$   | $\sigma_1 + \sigma_2$  |
| Series                                    | $\sigma_1.\sigma_2$ or $\sigma_1\sigma_2$                                 | $\sigma_1.\sigma_2$ or $\sigma_1\sigma_2$  |
| Parallel                                  | $\sigma_1 \sigma_2$   | $\sigma_1, \sigma_2$   |
|   | $P \circ Q$   | $\rho_1 \rho_2$<br>$P \circ Q$   |
| Replication                               | $!\sigma \doteq \sigma, \sigma, \dots$ (infinite)                         | $(n)\sigma \doteq \sigma, \sigma, \dots, \sigma; n$ parallel<br>$(\sigma)^n \doteq \sigma.\sigma.\dots.\sigma; n$ series |
|   | $!P \doteq P \circ P \circ \dots$ (infinite)                              | $(n)\rho \doteq \rho \rho \dots \rho; n$ parallel<br>$(n)P \doteq P \circ P \circ \dots \circ P; n$ parallel             |

while an Action is an element of the set,  $\sigma \in \mathcal{A}$  currently defined as:

$$\mathcal{A} \equiv \{ax; a \in \mathcal{S}, x \in \mathcal{C}\}, \quad (2)$$

where the set  $\mathcal{C}$  contains all possible channels. These notations are used throughout the text. In PABM, we use the following (minimal) set:

$$\mathcal{S}_{\min} \equiv \{m, b, 0, !\}; \quad (3)$$

where  $m$  is Mate,  $b$  is Bud, and two new actions,  $0$  and  $!$ , as the null and activate actions, respectively. The set  $\mathcal{A}$  remains the same but with  $\mathcal{S}$  replaced by  $\mathcal{S}_{\min}$ . We demonstrate that all other elements  $a \in \mathcal{S}$  (Eq. 1) can be derived from a combination of these modifications with the directed Actions of PBC. Note that with the changes, the action becomes a passive entity (i.e. an action waits for an activation signal) by default.

Since the definitions of mate and bud were not changed and have been discussed elsewhere [3, 5, 8], we will only review these definitions briefly.

**Activate action, !** Cardelli’s Brane calculus requires two levels of matching before an Action could be executed/activated: (a) input and output channels; and (b) action and co-action. The use of an activation signal is expected to

improve the symmetry of form for the Actions with respect to channel and activation pairings. Here, we introduce an activate action, ‘!’, to approximate the input/output channel functionality of stochastic pi calculus, consequently precluding the need for an explicit distinction between actions and coactions within Brane. The Action in the form  $!x$  acts as an initiator of membrane interaction through channel  $x$ . This Action may be interpreted as a binding event, analogous to the required output signal from the initiating membrane or molecule before any non-activate Action can be executed/activated.

The use of an activation signal, instead of a more strictly-bound action-coaction pair with matching channels, is based on the fact that a single compound, modeled here as a communication channel, can interact with more than one substance, which may range from proteins to oligosaccharides, on the cell membrane. As discussed in §2.2, typical biological interactions involving receptors, logically corresponding to channels, are one-to-many relationships, rather than one-to-one pairs. Nevertheless, such cardinality does not imply that reaction specificity is lost.

Another biological characteristic taken into account is the dependence of the kind of reaction that occurs on the receptor type, rather than on the ligand (i.e. it is the receiver that determines which effect will occur). This characteristic is particularly marked in cells of the immune system, as well as antibodies, which have different effector functions associated with each class and subclass. The proposed form emphasizes that interaction specificity is conferred by the channel, but the receiver determines the type of action to execute.

**The null action, 0.** The null action, 0, blocks actions that precede it; an Action in the form  $\sigma.0x_0$  can thus be used to model a blocked Action,  $\sigma$ . The null action can be deactivated with  $!x$ , making  $\sigma$  accessible. Biologically, blocking occurs in the event of temporary receptor internalization [9], binding-induced conformational changes [10, 11], and binding-induced physical blocking of other available binding sites. The use of 0 will be useful for modeling bind-and-release, molecular functionality switching, and other membrane-bound mechanisms.

**Bud.** Bud refers to the arbitrary splitting of a membrane, resulting in two membrane-bound compartments[3]. Cardelli makes a distinction between bud and *drip*; bud occurs when the split occurs with *one* internal membrane, while *drip* refers to the separation of *zero* internal membranes. In PABM, this distinction is not made.

**Mate.** Mate causes the irreversible mixing of actions of membranes that fuse either horizontally (i.e. membranes at the same level of nesting) or vertically through an exocytosis-type process[3].

## 2.2 Choice, Parallel, and Series

All discussions of choice, parallel and series compositions are made with reference to Actions, unless otherwise indicated. Parallel and series compositions are not valid for actions and channels.

**Choice.** The concepts of parallel composition, choice and prefix(series) are retained from pi calculus. The notation for choice will be retained ('+'). Choice could either be between actions  $a_1$  and  $a_2$  or channels  $x_1$  and  $x_2$ . These are equivalent to having a choice between two (or more) Actions in the basic form,  $ax$ . In particular, the following choices within action-channel pairs (Actions) would be equivalent to their respective choices between Actions:

$$(a_1 + a_2)x \equiv a_1x + a_2x \quad (4)$$

$$a(x_1 + x_2) \equiv ax_1 + ax_2 \quad (5)$$

$$(a_1 + a_2)(x_1 + x_2) \equiv a_1x_1 + a_2x_1 + a_1x_2 + a_2x_2 \quad (6)$$

These equivalences remove the need of implementing Actions in their non-basic forms. Hence, implementing choice in actions and/or channels will be unnecessary since all cases can always be reduced to a choice between (at least) two Actions.

Aside from simplifying the implementation of Actions, Eqns. 4 and 5 reflect biological phenomena. For instance, Eqn. 5 is illustrated by membrane-bound receptors that have multiple ligands, with each ligand binding with a different affinity. At least three virus families, *Orthomyxoviradae*, *Paramyxoviradae* and *Reoviradae*, for example, use sialic acid in cell surfaces to enter via the endocytic pathway(s)[14]. A biological phenomenon that illustrates Eqn. 4, on the other hand, is the receptor for advanced glycation of end products (RAGE), expressed in a wide variety of cell types. RAGE is characterized by its ability to recognize numerous ligands, each of which result in different effects[15]. This is equivalent to having several actions associated with the same channel. Although the reactions of RAGE do not involve membrane structure deformations, a feature that would allow the direct modeling of such events may be of interest. Furthermore, Eqn. 5 implies that several receptors (or channels), can be used to initiate the same actions. Different receptors, for instance, are used by different viruses to enter the cell. Equation 6 is included for the purpose of completeness, but may not have any biological significance.

**Parallel.** Parallel pi processes and Actions, as indicated in Table 1, are represented following the notations in Danos and Pradelier and Cardelli.

**Prefix/Series.** The original notation will be maintained for the series. A recurring series of the same Action would be used instead of replication to indicate the finite reusability of an Action.

### 2.3 Rates

PABM incorporates rates by associating a real number,  $r_x$ , to the channel of each Action  $ax$ . When  $r_x$  is associated with  $!x$ , it corresponds to the rate with which  $!x$  reacts on average with its receiver — the *basic rate*. When this real number is instead associated with  $ax$ ,  $a \neq !$ ,  $r_x$  is a factor of the basic rate which reflects the efficiency of the reaction. A value of 1.0 indicates that the specific reaction rate with a particular receiver is the same as the basic rate. Association of rates to channels is adapted from beta binders[4].

## 2.4 Affinity

Affinity describes the strength of non-covalent interactions between a ligand to its specific binding site on the receptor surface; this value is independent of the number of binding sites[12]. Higher affinities are associated with factors such as the exposure of large, interactive amino acid side-chains, highly electronegative groups, or the deformability of a surface; these characteristics generally enable a ligand to form more non-covalent bonds with the receptor[13]. Empirically, a value known as the affinity constant ( $K_a$ ) is used to approximate affinity for ligand-receptor systems<sup>1</sup>. It is determined by measuring the concentration of free ligand required to fill half of the binding sites on the receptor. When half the sites are filled,  $[\text{Ligand} \cdot \text{Receptor}] = [\text{Ligand}]$  and  $K_a = 1/[\text{Receptor}]$ , where ‘ $[X]$ ’ is used to indicate the molar concentration of  $X$ ; common  $K_a$  values range from as low as  $5 \times 10^4$  to as high as  $10^{11}$  liters/mole[12].

The use of affinity in process calculi for biology has been proposed by C. Priami and P. Quaglia as a feature for beta binders[4]. Affinity is incorporated as a probability  $P(a, b)$  that an interaction between two different interfaces  $a$  and  $b$  can take place, effectively relaxing the requirement for an exact matching of interface[6] — a distinct digression from pi calculus, where interactions occur on syntactically identical ports (lock-and-key model).

In PABM, affinity is inherent with choice. Since channels in PABM represent receptor-ligand functionality, the execution of a single action  $a$  can be associated with its interaction through more than one channel, say  $x_1, x_2, \dots$ , and  $x_n$ ,  $n > 1$ , resulting to the Action:  $a(x_1 + x_2 + \dots + x_n)$  that reduces to  $ax_1 + ax_2 + \dots + ax_n$  (Eqn. 5). When  $a = !$ , this results in multiple rates of execution, which depends on the Action  $a'x_l$  that is activated ( $a' \leq !$  and  $l \leq n$ ). A similar situation occurs when  $a \neq !$  and  $a' = !$ , albeit with a different biological implication. Table 2 summarizes the difference between this approach and Priami’s implementation in beta binders.

**Table 2.** Comparison of affinity in beta binders and PABM

|                | <i>Beta Binders</i>   | <i>PABM</i>                    |
|----------------|---|--------------------------------|
| Association    | each reduction  | each channel                   |
| Implementation | reaction probability $P(a, b)$ between two non-identical interfaces $a$ and $b$ | multiple channels using choice |

## 2.5 Branes and Systems

The definition for Systems as sets of nested Branes is retained, and notations for these are adapted from PBC[5]. Null Systems are represented as  $\diamond$ . Notations for parallel composition of Systems are also retained. The same replication rules are

<sup>1</sup> Notably for antibodies and antigens.

applied to both Actions and Systems (Table 1). Branes however, are redefined as a composition of Brane domains; a Brane consisting of a single domain reduces to the original definition (Table 1).

**Directed Brane domains and directed Actions.** In this section, the concept of directed actions in PBC is extended to *Brane domains*. A Brane domain, represented as a vector,  $\rho$ , is a grouping of directed Actions that approximates the occurrence of composition and functional non-homogeneity (“patchiness”) observed in biological membranes. Consequently, a Brane is now defined as a parallel composition of Brane domains. A Brane defined using a single domain is homogenous, and reduces to a Brane in PBC.

Brane domains were introduced to facilitate greater control in processes like membrane budding. As opposed to Cardelli’s calculus where a parameter is used to define the characteristics of the Brane that will be budded out, the proposed calculus makes these characteristics entirely dependent on the current, dynamic state of the parent membrane. Budding processes, however, are highly localized, and the derived system should *not* have all the characteristics of the parent membrane. In the proposed calculus, only specific Brane domains are transferred in budding processes, unless the parent membrane is homogenous.

Alternately, a Brane domain can be visualized as a set of directed Actions occurring proximally in a membrane. As an example, a system with Brane domains is subsequently represented as follows:

$$[\rho_1 | \rho_2 ] ( [ \rho_3 ] ( P ) \circ Q ) \quad (7)$$

where  $\rho_n$  is of the form  $\langle \sigma_1 ; \sigma_2 \rangle$  (Table 1). As in the original Cardelli calculus, both  $\rho_1$  and  $\rho_2$  are visible to  $\rho_3$ . Using the rules of PBC, only the “outside” Actions of  $\rho_3$  can interact with the “outside” Actions of both  $\rho_1$  and  $\rho_2$ . The advantage of this feature is relevant in modeling competition between parallel membrane processes (§3.6).

It is important to note that Brane domains represent active or functional sites on membranes or molecules and not the molecules themselves. Nevertheless, since at least one active site is associated with proteins, these can be represented as Brane domains. Brane domains can be used to model membrane proteins that function together such as lipid rafts[17] and SNARE complexes[18].

**Lateral and cross-membrane interaction.** Since interactions are now between two Brane domains, apart from the interaction of a domain from one membrane with another in a different membrane, domain-domain interactions within a membrane is now possible. Actions can now be activated by Activate Actions on a neighboring domain (lateral membrane interaction). Activations by Actions on opposite sides of a single membrane (cross-membrane) can also be facilitated, provided that one of the Actions is translocated to the other side of the membrane by a mechanism similar to diffusion or channel-mediated transport. This capability can be used to model ligands that interact with receptors

on the same membrane surface or on the opposite side of the same membrane. Spontaneous membrane and in-membrane operations such as pinocytosis, drip, inversion, and fusion of proteins to form rafts and complexes [17, 18] can now be easily modeled.

The following equation shows the competition between  $a_1x$  and  $a_2x$  since lateral- and cross-membrane interactions are allowed.

$$[ \langle !x ; a_1x \rangle | \langle a_2x ; - \rangle | \langle - ; a_3x \rangle ] ( \cdots ) \Rightarrow a_1 \text{ or } a_2 \quad (8)$$

Note that  $a_3$  cannot be activated since cross-membrane interaction are allowed only within the same Brane domain. With PABM, the transport of functional particles (e.g. molecules) through the membrane without introduction of atonal reduction rules can also be modeled (see §3.7).

**Volume information.** A single enzyme-substrate experiment in a controlled nanoenvironment has shown that the frequency of collisions between two molecules is inversely proportional to the size of the vesicle where these molecules are contained [19]. Consequently, volume information will be associated with each System, representative of a compartment, allowing adjustments to be made in the probabilities at which the contained reactions will occur.

## 2.6 Replication

For the purpose of a calculus geared towards discrete biological system modeling, PABM uses a more controlled form of replication for Actions (also applicable to Brane domains and Systems), where the cardinality of replication is indicated (see Table 1). For instance, even if the initial counts of cellular components that are in the order of  $\gtrsim 10^4$  to  $\sim 10^{10}$  [16] are large enough to warrant the use of  $\infty$ , these are still finite quantities that may be critical determinants of biological system viability, especially in simulations that run for relatively prolonged periods of time ( $\geq 24$  hours). Finite replication also reflects the finite lifetimes, masses and/or energies of both the components of biological systems and the systems themselves, appropriately manifested in the calculus in the form of finite Brane or Action usage. The numbers representing the finite number of replications can also be made stochastic to mimic the heterogeneity of membrane domains in terms of the absolute numbers of its constituents. Finite replication is conceptually similar to energy in beta binders [4], since the special entity  $E^j$  (with  $j \in \mathfrak{R}^+$ ) can be mapped to the cardinality of replication in PABM.

## 2.7 Sample Notation: Mitogen-Induced Proliferation of Schwann Cells

Cell proliferation induced by an external signal is one of the simplest biological examples that is, nevertheless, difficult to express in an intuitive manner without the use of spatial information. Mitogens, which induce cell division, are typically associated with one or more cognate receptors through which it can enter a cell. In Schwann cells, which form the insulation for vertebrate neurons

in the peripheral nervous system and which are critical for axon regeneration, proliferation is induced by the following mitogens in the neonatal stage: glial growth factor (GGF), platelet-derived growth-factor B (PDGF-BB) and basic fibroblast growth factor (bFGF) [32].

For purposes of illustration, a coarse-grained model of the system can be defined in the above notational changes as follows:

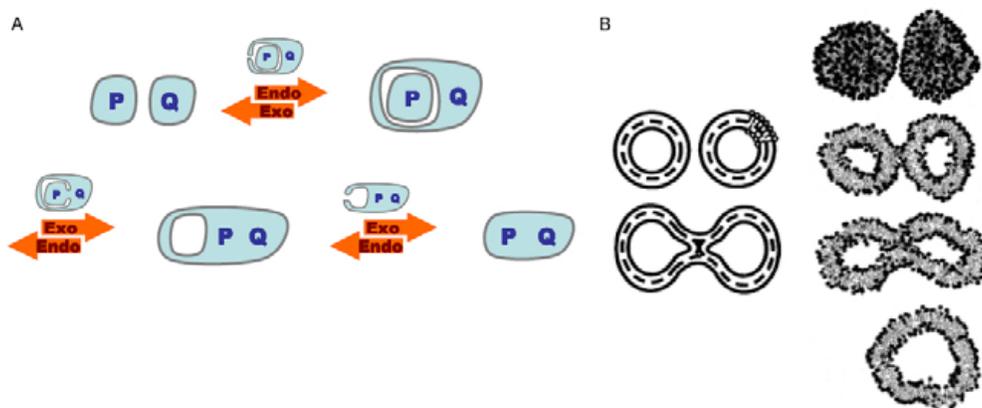
$$\begin{aligned} & [ \langle !x_G ; !x_a \rangle ] [ X ] \circ [ \langle !x_P ; !x_a \rangle ] [ X ] \circ [ \langle !x_{bF} ; !x_a \rangle ] [ X ] \circ \\ & [ \langle mx_G, mx_P, mx_{bF} ; bx_b \rangle ] [ SC \circ [ \langle !x_b \cdot 0x_a ; - \rangle ] [ R ] ] \end{aligned}$$

where each  $X$  represents a growth factor associated with some channel  $!x_s$ , where  $s$  represents the part of  $X$  that binds to the GGF receptor ( $G$ ), the PDGF-BB receptor ( $P$ ) or the bFGF receptor ( $bF$ ). The corresponding receptors,  $mx_G, mx_P, mx_{bF}$  are all associated with the Schwann cell ( $SC$ ).  $R$  represents the inactive replication machinery of the cell. This can only be activated on the fusion of one of the growth factors with  $SC$ , removing  $0x_a$ , and making  $!x_b$  available for interaction. The availability of  $!x_b$  in  $R$  allows  $SC$  to bud through its interaction with Action  $bx_b$ .

### 3 Projective Activate, Bud, and Mate Calculus

In this section, we demonstrate that all Actions in  $\mathcal{S}$  (Eq. 1) can be expressed as the actions in  $\mathcal{S}_{\min}$  (Eq. 3) combined with the directed Actions of PBC. The use of  $\mathcal{S}_{\min}$  as primitives has similarities to the basic Mate-Bud-Drip (MBD) calculus [8], which is one of two possible basic calculi for membrane interactions, together with the Phago-Exo-Pino (PEP) calculus. It has been shown [3, 8], however, that an encoding of MBD can be obtained with PEP, but not the opposite, because the maximum level of membranes (i.e. the membrane nesting) cannot grow during computation in MBD. Furthermore, the same articles prove that PEP calculus is Turing complete and Turing powerful, as opposed to MBD.

Given these limitations, the use of a Bud- and Mate-based basic calculus appears counterintuitive. However, events indicated in the derivation of the MBD primitives using PEP (Fig. 1A) are not observed in biological systems (Fig. 1B). Although it is partly superfluous to observe that the derivations of MBD were previously qualified as performed for computational purposes only, it is clear that *in vivo* membrane fusion is characterized by membrane perturbances rather than a series of phagocytosis and exocytosis events[20]. Specifically, the prevalent hypothesis regarding membrane fusion involves the reduction of the distance between the fusing membranes, followed by the local perturbation of the lipid structure and merger of proximal monolayers. Stalk formation and stalk expansion, and finally, pore formation are postulated to follow. Furthermore, there is a requirement that each of these steps has to be driven by an energy gradient towards lower energies. The stalk hypothesis is mainly based on the observation that the merger of proximal monolayers precedes the merger of distal monolayers. These events are followed by the intravesicular solvent exchange[20].



**Fig. 1.** PEP derivation of Mate [3] (A) and the latest model of how membrane fusion occurs [20] (B)

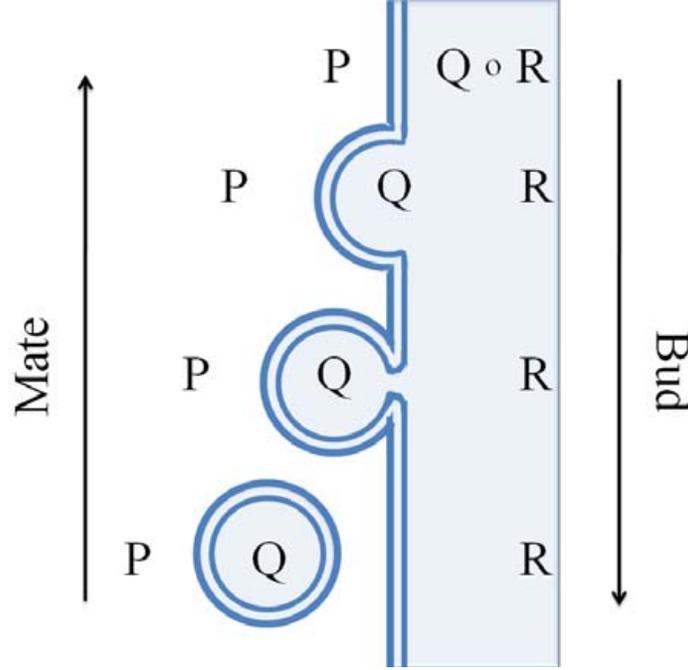
The succeeding discussions focus on the proposed PABM calculus as an encompassing calculus that conforms better with biologically observed phenomena.

### 3.1 Mate and Bud as Inverses of the Other

We consider Bud and Mate membrane actions as the primitives of this calculus, together with the Activate action, which controls their execution. Figure 2(top to bottom) shows a local deformation of the membrane separating the spaces labeled as  $P$  and  $Q \circ R$  resulting from its interaction with  $Q$ . The increase in local curvature is then followed by the movement of  $Q$  towards the newly formed protrusion. On the fusion of the initial points of deformation, a new membrane-bound space containing  $Q$  is formed within  $P$  (Fig. 2, bottom), completing Bud. The reverse process, Mate, can be obtained using an opposite perspective. Here, the membrane separating  $Q$  from  $P$  merges with the membrane separating  $P$  and  $R$  (bottom to top). Colors are used to indicate tonality; in these processes, bitonality is conserved, as in PBC[5]. Since Bud and Mate are opposite operations, it would be possible to think of these as belonging to a single operation.

### 3.2 Projective Equivalence

Projective equivalence arose from the introduction of directed actions by Danos and Pradalier [5]. Briefly, projective equivalence refers to the idea that the nature of membrane interactions is such that one does not make a distinction between top and bottom, or in this case, outside and inside. Consequently, by using a simple point-of-view change (i.e. what one considers inside before, which is a bounded space, is now viewed as the outside, which is unbounded), one reverses the process. If one uses a pointed bitonal tree representation for the structure, the equivalence is simply a change in the *distinguished vertex*[5], which is a change in the root of the tree. One can then generalize phago and bud as a single budding action, and exo and mate as a single mate action.



**Fig. 2.** Mate and Bud as inverse actions of the other. Bud is shown as a sequence from top to bottom while Mate as the reverse. Note that a distinction is not made between “inside” and “outside” spaces. A bilayer is used to illustrate directionality.

### 3.3 Basic Reduction Rules

The basic reduction rules of PABM are entirely based on Bud and Mate. Reduction rules are applied between interacting Brane domains, where the location of the activation signal with respect to the receiver (i.e. the directionality of the Action) determines if a Bud will be inward or outward, or if a Mate will be horizontal (i.e. membranes at the same level will merge) or vertical (i.e. the membrane of a content will merge with the membrane of its parent); this is conceptually similar to what has been done in PBC[5]. In the design of a Brane model, the directions at which Bud and Mate proceed are naturally integrated. The reduction rules of PABM are as follows, with “ $\sim$ ” used for indicating projective equivalence[5].

– Bud:

$$\begin{aligned} & P \circ [\rho_1 | \langle \sigma_1 ; \sigma_2, \tau_2 . bx \rangle ] \llbracket [\rho_2 | \langle \sigma_4, \tau_4 . !x ; \sigma_3 \rangle ] \llbracket Q' \rrbracket \circ R \rrbracket \\ \longrightarrow & P \circ [ \langle \sigma_1 ; \sigma_2, \tau_2, \tau_4 \rangle ] \llbracket [\rho_2 | \langle \sigma_4 ; \sigma_3 \rangle ] \llbracket Q' \rrbracket \rrbracket \circ [ \rho_1 ] \llbracket R \rrbracket \end{aligned} \quad (9)$$

$$\begin{aligned} & \llbracket [\rho_1^\dagger | \langle \sigma_2, \tau_2 . bx ; \sigma_1 \rangle ] \llbracket P \rrbracket \rrbracket \circ [ \rho_2 | \langle \sigma_4, \tau_4 . !x ; \sigma_3 \rangle ] \llbracket Q' \rrbracket \circ R \\ \longrightarrow & \llbracket [\rho_1^\dagger ] \llbracket P \circ [ \langle \sigma_1 ; \sigma_2, \tau_2, \tau_4 \rangle ] \llbracket [\rho_2 | \langle \sigma_4 ; \sigma_3 \rangle ] \llbracket Q' \rrbracket \rrbracket \rrbracket \rrbracket \circ R \end{aligned} \quad (10)$$

– Mate:

$$\begin{aligned} & P \circ [\rho_2 | \langle \sigma_1, \tau_1. !x ; \sigma_2 \rangle ] ( Q ) \circ [\rho_1 | \langle \sigma'_1, \tau_3. mx ; \sigma'_2 \rangle ] ( R ) \\ \longrightarrow & P \circ [\rho_1 | \rho_2 | \langle \sigma_1 ; \sigma_2 \rangle | \langle \sigma'_1, \tau_1, \tau_3 ; \sigma'_2 \rangle ] ( Q \circ R ) \end{aligned} \quad (11)$$

$$\begin{aligned} & \left[ \rho_1^\dagger | \langle \sigma'_2 ; \sigma'_1, \tau_3. mx \rangle \right] ( P \circ [\rho_2 | \langle \sigma_1, \tau_1. !x ; \sigma_2 \rangle ] ( Q ) ) \circ R \\ \longrightarrow & \left[ \rho_1^\dagger | \rho_2 | \langle \sigma_2 ; \sigma_1 \rangle | \langle \sigma'_1, \tau_1, \tau_3 ; \sigma'_2 \rangle \right] ( P ) \circ Q \circ R \end{aligned} \quad (12)$$

Note that  $Q$  in Mate is equivalent to  $[\sigma_4, \tau_4. !x ] (\sigma_3 ) Q'$  in Bud. Odd-even subscripts and primed Actions are used to illustrate bitonality preservation. For directionality to be conserved, note the need for the reversal of  $\rho_1$  to  $\rho_1^\dagger$  when the perspective is changed.

In the case of Mate, interchanging the locations of  $mx$  and  $!x$  results in slightly different Brane domains. Eq. 11 results in:

$$P \circ [\rho_1 | \rho_2 | \langle \sigma_1, \tau_1, \tau_3 ; \sigma_2 \rangle | \langle \sigma'_1 ; \sigma'_2 \rangle ] ( Q \circ R ) ; \quad (13)$$

while Eq. 12 results in:

$$[\rho_1 | \rho_2 | \langle \sigma_1 ; \sigma_2 \rangle | \langle \sigma'_1, \tau_1, \tau_3 ; \sigma'_2 \rangle ] ( P ) \circ Q \circ R. \quad (14)$$

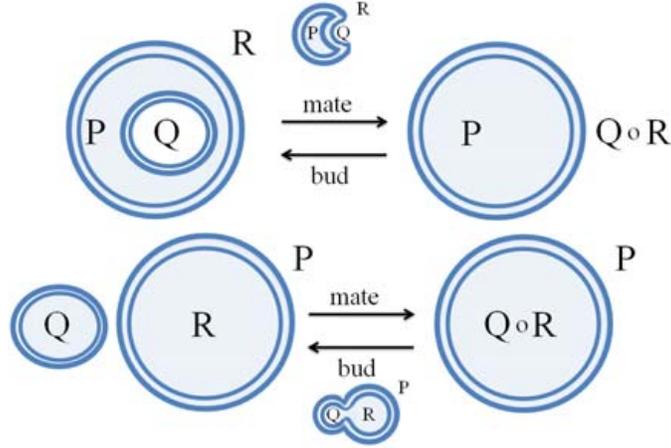
It is only in the absence of  $\tau_1$  and  $\tau_3$  that the location of  $mx$  and  $!x$  does not result in different succeeding states.

### 3.4 Non-primitive Actions with Bud and Mate

As shown in Eqs. 9 to 12, congruence exists between an inward and outward Bud, and between a horizontal and vertical Mate. This is more clearly illustrated in Fig. 2, where one sees that a simple perspective shift makes the same Bud or Mate operation inward or outward, or vertical or horizontal. For instance, when one chooses  $P$  as the “inside”,  $Q$  can be viewed as budding in towards  $P$ , or that the membrane containing  $Q$  is mating with the membrane separating  $P$  and  $R$ . As a result, PABM considers  $a \in \mathcal{S}$ ,  $a \notin \mathcal{S}_{\min}$  as membrane operations congruent to either Bud or Mate operation or its specific cases.

**Phago and Exo.** Fig. 3(top) shows how  $Q$  is exocytosed from  $R$  or endocytosed into  $P$  via Mate and Bud, respectively. This is congruent to the Mate-Bud reactions in Fig. 2, with  $R$  as the outside and  $P$  as the inside. Phago is expressed as Bud in Eq. 10 and is congruent to the usual bud in Eq. 9.

**Pino and Drip.** Pino and Drip may be spontaneous or induced Bud actions (see Eq. 10), where a null System ( $Q = \text{null}$ ) is created inside or outside the bounded space  $P$ . The activation may also be induced by an appropriate Activate Action outside or inside  $P$ , or within-membrane activations (see §2.5). Pino and drip are obtained when  $Q = \text{null}$  in Fig. 3.



**Fig. 3.** Specific cases of Mate and Bud: (top, forward) Exo and (top, reverse) Phago; (bottom, forward) Cardelli’s Mate and (bottom, reverse) Bud operations

### 3.5 Enhanced Membrane Dynamics

A fundamental difference of the proposed calculus from the Brane calculus of Cardelli is the dynamic nature of the reacting membranes. In Cardelli’s version, the properties of budded membranes are specified as parameters to provide control; the same is true for the “endosomes” formed during Phago. In biological systems, however, the characteristics of the budded membrane are necessarily dependent on the *state* of the parent membrane at the time of budding or phagocytosis. Fig. 4 reflects this particular case of budding, when a sequential action is associated with the activation Action on an initiating membrane ( $\sigma_1$ , green). Note the incorporation of  $\sigma_1$  in the budded membrane. Also note that only a portion of the membrane is budded out.

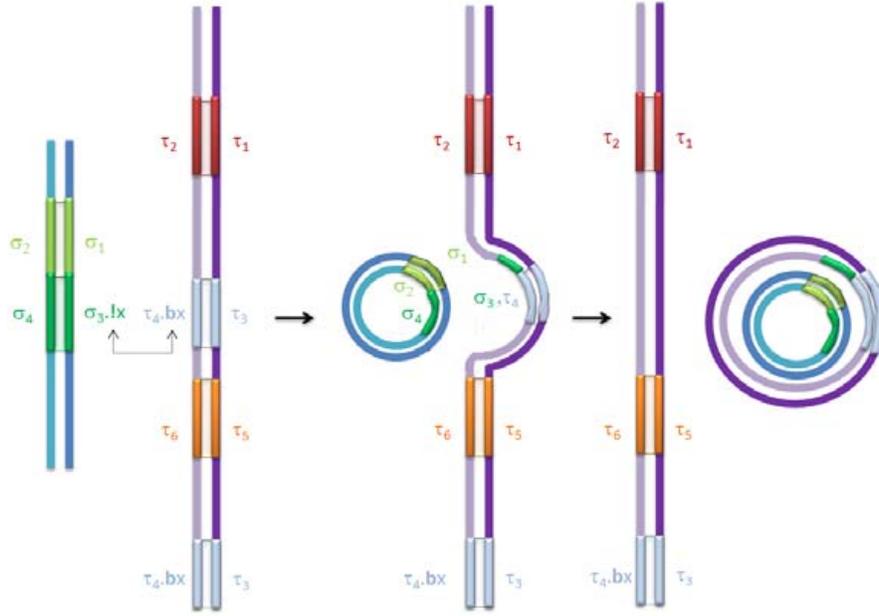
This dynamic property of the membrane generally implies that systems involved in a Mate followed by a Bud ( $b.m$ ) will not evolve equivalently when Bud is performed before Mate ( $m.b$ ). For instance, consider the following initial system:

$$Q_0 \equiv [ \langle \sigma_4, \tau_4.m x_M ; \sigma_3, \tau_3.b x_B \rangle ] [ \langle \sigma_1, \tau_1.!x_B ; \sigma_2 \rangle ] [ P_2 ] \circ P_3 ] \circ [ \langle \sigma_6, \tau_6.!x_M ; \sigma_7 \rangle ] [ P_1 ] ] \quad (15)$$

Depending on which operation occurs first, the system will evolve in two different ways. First, on performance of  $b.m$ , the system will evolve as:

$$Q_0 \xrightarrow{m} Q'_0 \equiv [ \langle \sigma_4, \tau_4, \tau_6 ; \sigma_3, \tau_3.b x_B \rangle ] [ \langle \sigma_6 ; \sigma_7 \rangle ] [ P_1 \circ [ \langle \sigma_1, \tau_1.!x_B ; \sigma_2 \rangle ] [ P_2 ] \circ P_3 ] ] \quad (16)$$

$$Q'_0 \xrightarrow{b} Q'_1 \equiv [ \langle \sigma_6 ; \sigma_7 \rangle ] [ P_1 \circ P_3 ] \circ [ \langle \sigma_4, \tau_4, \tau_6 ; \sigma_3, \tau_1, \tau_3 \rangle ] [ [ \langle \sigma_1 ; \sigma_2 \rangle ] [ P_2 ] ] ] \quad (17)$$



**Fig. 4.** Budding as a dynamic process. Note that a new action,  $\sigma_1$ , associated with the ‘!’ is incorporated into the budded membrane. Domains are illustrated as line segments; only selected domains proximal to the activated action are budded out.

Second, with  $m.b$ , the system will evolve as:

$$Q_0 \xrightarrow{b} Q''_0 \equiv [- \llbracket P_3 \rrbracket \circ [ \langle \sigma_4, \tau_4.m x_M ; \sigma_3, \tau_3, \tau_1 \rangle \llbracket [ \langle \sigma_1 ; \sigma_2 \rangle \llbracket P_2 \rrbracket \rrbracket \circ [ \langle \sigma_6, \tau_6.!x_M ; \sigma_7 \rangle \llbracket P_1 \rrbracket ] ] ] \quad (18)$$

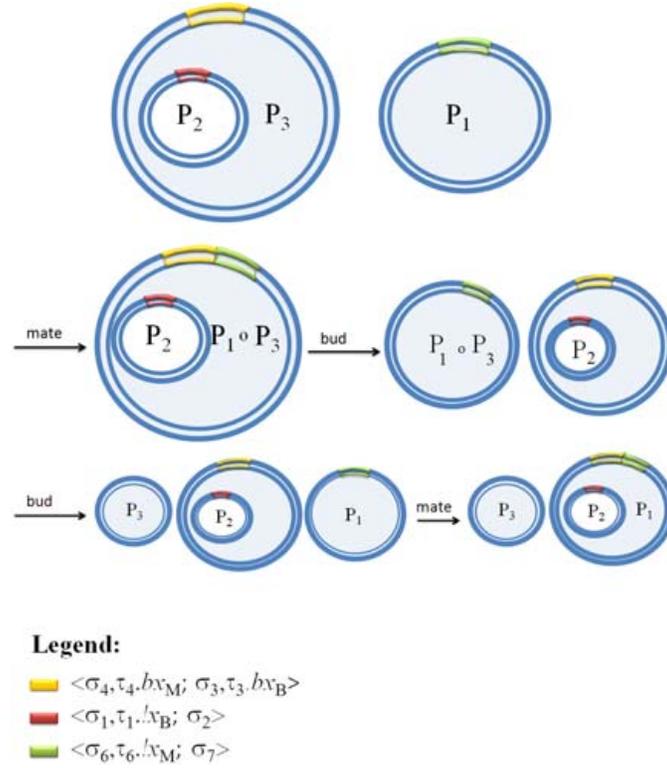
$$Q''_0 \xrightarrow{m} Q''_1 \equiv [- \llbracket P_3 \rrbracket \circ [ \langle \sigma_4, \tau_4, \tau_6 ; \sigma_3, \tau_3, \tau_1 \rangle \llbracket \langle \sigma_6 ; \sigma_7 \rangle \llbracket P_1 \circ [ \langle \sigma_1 ; \sigma_2 \rangle \llbracket P_2 \rrbracket \rrbracket ] ] ] \quad (19)$$

Clearly,  $Q'_1 \neq Q''_1$ . This asymmetry example ( $b.m \neq m.b$ ) is depicted in Fig. 5. The difference disappears when the Mate and Bud are placed in separate domains,  $\langle \sigma_4, \tau_4.m x_M ; \sigma_3 \rangle$  and  $\langle \sigma'_4 ; \sigma'_3, \tau_3.b x_B \rangle$ .

### 3.6 Competition of Parallel Membrane Processes

Using the concept of Brane domains, competition of two or more parallel membrane processes can be easily modeled. Given the following system (longhand), the interactions of  $!x$  is restricted to  $bx$  associated with  $\sigma_1$  or that associated with  $\sigma_3$ . If it interacts with  $bx$  in  $\langle \sigma_1 ; bx \rangle$ , then the system reduction will be in the form:

$$\begin{aligned} & [ \langle \sigma_1 ; bx \rangle \llbracket \langle \sigma_3 ; bx \rangle \llbracket [ \langle !x ; \sigma_5 \rangle \llbracket Q' \rrbracket \rrbracket \circ R ] ] \\ \longrightarrow & [ \langle \sigma_1 ; - \rangle \llbracket [ \langle - ; \sigma_5 \rangle \llbracket Q' \rrbracket \rrbracket \rrbracket \circ [ \langle \sigma_3 ; bx \rangle \llbracket R \rrbracket ] ] \end{aligned} \quad (20)$$



**Fig. 5.** Differences in final system states based on the order at which reactions occur. (top) Initial configuration; (middle) Mate then Bud; (bottom) Bud then Mate.

On the other hand, if it interacts with  $bx$  in  $\langle \sigma_3; bx \rangle$  instead, then the system reduction will be as follows:

$$\begin{aligned}
 & [ \langle \sigma_1; bx \rangle | \langle \sigma_3; bx \rangle ] | [ \langle !x; \sigma_5 \rangle ] | ( Q' ) \circ R ) \\
 \longrightarrow & [ \langle \sigma_3; - \rangle ] | [ \langle -; \sigma_5 \rangle ] | ( Q' ) \circ [ \langle \sigma_1; bx \rangle ] | R ) \quad (21)
 \end{aligned}$$

Competition can also be realized in lateral and cross-membrane processes (see Eq. 8).

### 3.7 Molecules as Systems

Molecules can either serve as ligands or receptors. In this proposed modification, Molecules can be modeled as null Systems containing Activators associated with Actions or other Activators. It may also be in the form of blocking functions,  $\sigma.0x$ , which could only be activated by  $!x$ . A molecule can be modeled as:

$$\text{Molecule} : [ \langle \sigma_1, (n)(!x_{\text{name}} + !x_{\text{generic}}); \sigma_2, (m)(!x_{\text{name}} + !x_{\text{generic}}) \rangle ] | ( \diamond ) \quad (22)$$

where  $\tau$  may be a null Brane and the “name” could be the name of the molecule making the channel unique for the molecule and “generic” refers to the generic channel of the molecule. For example, “RNA” can be a generic channel having the name of the protein that it encodes for its specific name.

Together with cross-membrane interactions, molecule diffusion through a membrane can be modeled without atonal reduction. This is illustrated in the reduction below, where the molecule on the left-hand side enters the system  $P$ . Note the change in the replication coefficient, reflecting the reduction of the active sites in both the molecule and the membrane surrounding  $P$ .

$$\begin{aligned}
& [ \langle (!x)^{(n_1)} ; (bx.!y)^{(n_1)} \rangle ] (| \diamond |) \circ [ (n_2) \langle my ; - \rangle ] (| P |) \\
\longrightarrow & [ (n_2 - 1) \langle my ; - \rangle ] \langle (bx.!y)^{(n_1-1)}.bx ; (!x)^{(n_1-1)}.!x \rangle (| P |) \\
\longrightarrow & [ (n_2 - 1) \langle my ; - \rangle ] (| P \circ [ \langle (!x)^{(n_1-1)} ; (bx.!y)^{(n_1-1)} \rangle ] (| \diamond |) |) \quad (23)
\end{aligned}$$

### 3.8 Mass and Energy Conservation

Since budding involves direct movement of Brane domains, mass (represented by an Action) conservation is also simulated. The “consumption” of an Action after reduction can be seen as the usage of the available energy used for and/or transfer of mass during the process, i.e. transformation of the structure into new ones.

## 4 PABM as an Extension of Existing Brane Calculi

Equivalent expressions for the multiple association of the activation action using Cardelli’s original notations can be derived. Suppose there are three systems that can interact via a generic action  $a$ , with coaction  $\bar{a}$ . Multiple association can be realized with:

$$a!x (| Q_0 |) \circ \bar{a}?x (| P_1 |) \circ \bar{a}?x (| P_2 |), \quad (24)$$

where system  $Q_0$  can proceed with the action ( $a \leftrightarrow \bar{a}$ ) on both systems  $P_1$  and  $P_2$  through the same channel  $!x \rightarrow ?x$ . It is possible to eliminate the use of coactions ( $\bar{a}$ ) through the following representation:

$$!x (| Q_0 |) \circ a?x (| P_1 |) \circ a?x (| P_2 |) \quad (25)$$

with  $!x$  possibly activating either  $P_1$  or  $P_2$  via  $a?x$ . This minor notation change is immediately compatible with Pi calculus, and would require a minor code translation for recognition by the Stochastic Pi Machine (SPiM) [21].

However, the proposed calculus also involves the removal of the sender-receiver pairing (viz.  $!x \rightarrow ?x$ ), apart from the action-coaction pairing. Moreover, in Eq. 25, the notation is asymmetric, with the activator  $!x$  having a different form from  $a?x$ . It is possible to use  $\alpha!x$  as a universal activator to conserve symmetry, but  $\alpha$  will be underutilized. The use of a single “sender” is proposed for all the other actions in the form  $!x$ , while *simultaneously* making the notation symmetric with the use of the same form ( $ax$ , see Table 1). Both conceptually and implementationwise, these major differences could be seen as improvements over the current representation.

Hence, the multiple association expressed as (24) would be written in PABM as:

$$[ \langle !x ; - \rangle \langle Q_0 \rangle ] \circ [ \langle ax ; - \rangle \langle P_1 \rangle ] \circ [ \langle ax ; - \rangle \langle P_2 \rangle ] \quad (26)$$

with  $x$  as the channel; “!” now belongs to the same class as  $a \neq !$  (Table 1). The choice of the symbol “!” for the activate action is directly inspired by the Pi calculus notation.

As indicated previously, the other major departure from Cardelli’s Brane calculus is the utilization of Brane domains in dynamic membranes to eliminate the use of parameters in phagocytosis and budding. With these domains, a Brane in PBC becomes a special case when a Brane in PABM is homogenous (i.e. is comprised of a single Brane domain). For purposes of comparison, the reduction rules of the original calculus [taken from [8]] are shown in Fig. 6.

$$\begin{array}{ll}
\text{(phago)} & \mathfrak{D}_n.\sigma|\sigma_0\langle P \rangle \circ \mathfrak{D}_n^\perp(\rho).\tau|\tau_0\langle Q \rangle \rightarrow \\
& \tau|\tau_0\langle \rho|\sigma|\sigma_0\langle P \rangle \rangle \circ Q \\
\text{(exo)} & \mathfrak{D}_n^\perp.\tau|\tau_0\langle \mathfrak{D}_n.\sigma|\sigma_0\langle P \rangle \circ Q \rangle \rightarrow \\
& P \circ \sigma|\sigma_0|\tau|\tau_0\langle Q \rangle \\
\text{(pino)} & \odot(\rho).\sigma|\sigma_0\langle P \rangle \rightarrow \sigma|\sigma_0\langle \rho \rangle \circ P \\
\text{(mate)} & \text{mate}_n.\sigma|\sigma_0\langle P \rangle \circ \text{mate}_n^\perp.\tau|\tau_0\langle Q \rangle \rightarrow \\
& \sigma|\sigma_0|\tau|\tau_0\langle P \circ Q \rangle \\
\text{(bud)} & \text{bud}_n^\perp(\rho).\tau|\tau_0\langle \text{bud}_n.\sigma|\sigma_0\langle P \rangle \circ Q \rangle \rightarrow \\
& \rho|\sigma|\sigma_0\langle P \rangle \rangle \circ \tau|\tau_0\langle Q \rangle \\
\text{(drip)} & \text{drip}(\rho).\sigma|\sigma_0\langle P \rangle \rightarrow \rho \langle \rangle \circ \sigma|\sigma_0\langle P \rangle
\end{array}$$

**Fig. 6.** Cardelli’s Brane calculus reduction rules taken from [8]

The realization of non-primitive actions that were illustrated utilizes the same concepts as in the projective equivalence of Danos and Pradalier [5], with the exception that no arguments are explicitly used for the Bud action. Furthermore, the simplicity of the current basis and reduction makes the calculus closer to actual biological membrane operations. Finally, PBC becomes a subset of the proposed calculus since PBC Branes can be simulated using homogenous PABM Branes.

## 5 Summary and Outlook

We end this paper with brief discussions on a potential application of the calculus, as well as a strategy for its possible implementation.

## 5.1 Application Example: Viral Infection

Influenza A causes highly contagious respiratory infections in humans that range in severity from acute to lethal. New strains arise annually, which lead to 250,000 to 500,000 deaths worldwide[22, 23, 24]. It is particularly interesting for biologists because of its ability to evolve very quickly, a trait that makes the development of an efficient vaccine against it particularly challenging[25, 26]. To date, a number of qualitative studies have been performed to investigate its life cycle, but most involved separate analyses of steps in the infection process [24, 26].

One of the recent most extensive quantitative models of influenza A in cell culture is that by Sidorenko and Reichl[24]. It consists of 49 ordinary differential equations (ODEs) that involve the use of additional parameters to approximate the movement of viruses and its components across cell compartments. The main results obtained from the model include the identification of factors that limit the growth rate of viral progeny; these results are particularly useful in molecular engineering, where engineered viruses are created for vaccine production[24]. Nevertheless, it is clear that much is still not known about the influenza A life cycle, primarily owing to the complexity of the virus. Some details, for instance, that have not been included in the Sidorenko-Reichl model include the following:

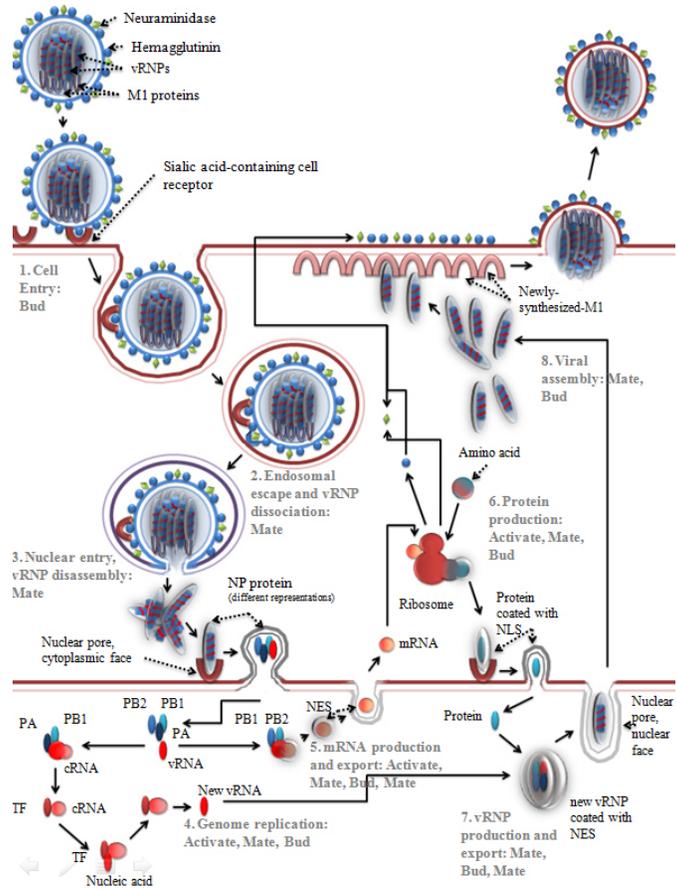
1. Distinction between each of the eight strands of genetic material (vRNA), complexed with three proteins (collectively known as vRNP), throughout the replication cycle
2. Distribution of 11 protein-coding genes across the eight vRNAs
3. Indirect genetic material replication (vRNA  $\rightarrow$  cRNA  $\rightarrow$  vRNA), with the intermediate cRNA being able to interact with the same proteins that vRNA interacts with
4. Requirement for precise viral assembly

Accordingly, several key issues remain unanswered:

1. time it takes to assemble vRNPs
2. ratio of infective to non-infective viruses
3. instances of ‘infectivity recovery’ in the event that two complementary non-infective viruses enter a cell

Since compartments can be naturally represented in Brane, its use for modeling the influenza A life cycle is probably an elegant, quantitative alternative that would allow the inclusion of details such as those enumerated previously. Fig. 7 is a general illustration for the possible usage of PABM to model the influenza infection cycle. This particular model is an interesting application for Brane calculus on account of its scale. Note that all operations used for modeling the system are restricted to budding and mating, including the simulation of the bind-and-release action in the nucleus. The position of activation signals are not explicitly indicated in the figure, but could be deduced from the illustrated processes.

In addition to these, it would also be possible to include details that are known to affect influenza infectivity, as well as efficiency[27]:



**Fig. 7.** Influenza A infection cycle model being implemented using PABM. All operations are performed with Bud and Mate, including nuclear import and export through the use of the NP protein, nuclear localization sequences (NLS) or nuclear export sequences (NES). All processes are conformant with the actual events in influenza infection.

1. cleavage efficiency of HA
2. distinction between transcriptionally active and inactive vRNPs

## 5.2 Implementation and Compatibility with SPiM

Previous efforts have been made to implement Brane calculus[28, 29]. These implement calculi based on the set  $\mathcal{S}$  (Eq. 1) and were found useful for studying events having the same scale as the Semliki forest virus life cycle, which was used as the illustrative example in [3]. Nevertheless, these are not powerful enough to handle models having the scale of the influenza A life cycle. It is consequently of interest to develop an implementation that is both scalable and robust.

The stochastic pi machine (SPiM) was developed by Andrew Phillips, and uses a simulation algorithm for stochastic pi calculus that is particularly suited for simulating biological systems involving a large number of molecules. This

simulation algorithm makes the execution cost dependent on the number of species, rather than the actual number of molecules, unlike in direct implementations of the Gillespie algorithm [21]. SPiM has been used on a number of occasions for a variety of biological problems [21, 30, 31]. Lately, the algorithm in SPiM has been extended to include compartment-based computation, using the Bioambients formalism [7]. SPiM also has a graphical interface, which significantly improves its ease of use.

PABM should be compatible with SPiM using the following equivalences:

$$!x \equiv (m!x + b!x + 0!x) \quad (27)$$

$$ax \equiv \bar{a}?x \quad (28)$$

$$\sigma.!x \equiv a!x(\sigma) \quad (29)$$

where  $a = m, b, 0$  and  $\bar{a}$  is the corresponding coaction. Encoding more specific stochastic pi calculus constructs would only require the use of very specific channel names. For PABM to be implemented on top of SPiM, compartments, Brane domains, and action directionality have to be appropriately represented. A separate implementation approach that focuses on the rewriting rules of PABM is also currently being explored.

*Acknowledgments.* We wish to thank Luca Cardelli and Andrew Phillips for helpful discussions.

## References

- [1] Păun, G.: Introduction to membrane computing. In: Applications of Membrane Computing, pp. 1–42 (2006)
- [2] Regev, A., Shapiro, E.: The  $\pi$ -calculus as an abstraction for biomolecular systems. In: Modelling in Molecular Biology (2004)
- [3] Cardelli, L.: Brane calculi: interactions of biological membranes. In: Danos, V., Schachter, V. (eds.) CMSB 2004. LNCS (LNBI), vol. 3082, pp. 257–278. Springer, Heidelberg (2005)
- [4] Priami, C., Quaglia, P.: Beta binders for biological interactions. In: Danos, V., Schachter, V. (eds.) CMSB 2004. LNCS (LNBI), vol. 3082, pp. 20–33. Springer, Heidelberg (2005)
- [5] Danos, V., Pradalier, S.: Projective brane calculus. In: Danos, V., Schachter, V. (eds.) CMSB 2004. LNCS (LNBI), vol. 3082, pp. 134–148. Springer, Heidelberg (2005)
- [6] Guerriero, M.L., Priami, C., Romanel, A.: Modeling Static Biological Compartments with Beta-binders. In: Anai, H., Horimoto, K., Kutsia, T. (eds.) Ab 2007. LNCS, vol. 4545, pp. 247–261. Springer, Heidelberg (2007)
- [7] Phillips, A., Cardelli, L.: Efficient, correct simulation of biological processes in the stochastic pi-calculus. In: Calder, M., Gilmore, S. (eds.) CMSB 2007. LNCS (LNBI), vol. 4695, pp. 184–199. Springer, Heidelberg (2007)
- [8] Busi, N., Zandron, C.: Modelling and analysis of biological processes by (mem)brane calculi and systems. In: Proceedings of the 2006 Winter Simulation Conference, pp. 1646–1655 (2006)

- [9] Guglielmo, G.D., Drake, P., Baass, P., Authier, F., Posner, B., Bergeron, J.: Insulin receptor internalization and signalling. *Mol. Cell. Biochem.* 182, 59–63 (1998)
- [10] Hsu, S., Bonvin, A.: Atomic insight into the CD4 binding-induced conformational changes in HIV-1 gp120. *Proteins: structure, function and bioinformatics* 3, 582–593 (2004)
- [11] Keskin, O.: Binding induced conformational changes of proteins correlate with their intrinsic fluctuations: a case study of antibodies. *BMC Structural Biology* 7, 31 (2007)
- [12] Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., Watson, J.: *Molecular Biology of the Cell*, New York (2002)
- [13] David, M., Asprer, J., Ibane, J., Concepcion, G., Padlan, E.: A study of the structural correlates of affinity maturation: antibody affinity as a function of chemical interactions, structural plasticity and stability. *Mol. Immunol.* 44, 1342–1351 (2006)
- [14] Dimitrov, D.: Virus entry: molecular mechanisms and biomedical applications. *Nature Reviews Microbiology* 2, 109–122 (2004)
- [15] Kim, W., Hudson, B., Moser, B., Guo, J., Rong, L., Yu, L., Qu, W., Lalla, E., Lerner, S., Chen, Y., Yan, S.D., D'Agati, V., Naka, Y., Ramasamy, R., Herold, K., Yan, S., Schmidt, A.: Receptor for advanced glycation end products and its ligands: A journey from the complications of diabetes to its pathogenesis. *Annals of the New York Academy of Sciences* 1043, 553–561 (2006)
- [16] Thulke, S., Radonic, A., Nitsche, A., Siegert, W.: Quantitative expression analysis of HHV-6 cell receptor CD46 on cells of human cord blood, peripheral blood and G-CSF mobilised leukapheresis cells. *Virology Journal* 3, 77–81 (2006)
- [17] Simons, K., Vaz, W.L.C.: Model Systems, Lipid Rafts, and Cell Membranes. *Annual Review of Biophysics and Biomolecular Structure* 33(1) (June 2004)
- [18] Sutton, R.B., Fasshauer, D., Jahn, R., Brunger, A.T.: Crystal structure of a SNARE complex involved in synaptic exocytosis at 2.4 Å resolution. *Nature* 395(6700), 347–353 (1998)
- [19] Chiu, D., Wilson, C., Karlsson, A., Danielsson, A., Lunqvist, A., Stroemberg, A., Ryttsen, F., Davidson, M., Nordholm, S., Orwar, O., Zare, R.: Manipulating the biochemical nanoenvironment around single molecules contained within vesicles. *Chem. Phys.* 247, 133–139 (1999)
- [20] Jahn, R., Grubmüller, H.: Membrane fusion. *Current Opinion in Cell Biology* 14, 488–495 (2002)
- [21] Phillips, A., Cardelli, L.: A Correct Abstract Machine for the Stochastic Pi-calculus. In: *Concurrent Models in Molecular Biology* (2004)
- [22] Poland, G.A., Tosh, P., Jacobson, R.M.: Requiring influenza vaccination for health care workers: seven truths we must accept. *Vaccine* 23(17-18), 2251–2255 (2005); *Vaccines and Immunisation. Based on the Fourth World Congress on Vaccines and Immunisation*
- [23] Baccam, P., Beauchemin, C., Macken, C.A., Hayden, F.G., Perelson, A.S.: Kinetics of Influenza A Virus Infection in Humans. *J. Virol.* 80(15), 7590–7599 (2006)
- [24] Sidorenko, Y., Reichl, U.: Structured model of influenza virus replication in mdck cells. *Biotechnology and bioengineering* 88, 1–14 (2004)
- [25] Bardiya, N., Bae, J.: Influenza vaccines: recent advances in production technologies. *Applied Microbiology and Biotechnology* 67(3), 299–305 (2005)
- [26] Genzel, Y., Schulze-Horsel, J., Möhler, L., Sidorenko, Y., Reichl, U.: Influenza vaccines –challenges in mammalian cell culture technology. *Cell Technology for Cell Products*, 503–508 (2007)

- [27] Nayak, D.P., Hui, E.K.-W., Barman, S.: Assembly and budding of influenza virus. *Virus Research* 106, 147–165 (2004)
- [28] de Ronde, J.J., Ndjehan, C.P.: Modelling Networks and Pathways in Systems Biology. Technical report, CA545 Practicum, School of Computing, Dublin City University (2005/2006)
- [29] David, M.P.C.: BCD: Design and implementation of a stochastic brane machine. Master's thesis, Department of Computer Science, University of the Philippines, Diliman, Quezon City (2008)
- [30] Segata, N., Blanzieri, E., Priami, C.: Stochastic  $\pi$ -calculus modelling of multisite phosphorylation based signaling: in silico analysis of the Pho4 transcription factor and the PHO pathway in *Saccharomyces cerevisiae*. Technical report, Center for Computational and Systems Biology, The Microsoft Research – University of Trento (2007)
- [31] Yap, J.M.: A Pi-Calculus Model of the CD95 Receptor Medicated Pathway of Apoptosis. *Philippine Information Technology Journal* 1(1) (2008)
- [32] Zhang, B.T., Hikawa, N., Horie, H., Takenaka, T.: Mitogen induced proliferation of isolated adult mouse Schwann cells. *J. Neurosci. Res.*, 648–654 (1995)



# Appendix C

## Original publication P3

Maria Pamela Dobay, Alicia Piera Alberola, Eduardo Mendoza, Joachim O. Rädler  
**Modeling nanoparticle uptake and intracellular distribution using stochastic process algebras**  
*J Nanopart Res* **14**: 821 - 833, 2012.



# Modeling nanoparticle uptake and intracellular distribution using stochastic process algebras

M. P. D. Dobay · A. Piera Alberola ·  
E. R. Mendoza · J. O. Rädler

Received: 20 September 2011 / Accepted: 10 March 2012  
© Springer Science+Business Media B.V. 2012

**Abstract** Computational modeling is increasingly important to help understand the interaction and movement of nanoparticles (NPs) within living cells, and to come to terms with the wealth of data that microscopy imaging yields. A quantitative description of the spatio-temporal distribution of NPs inside cells; however, it is challenging due to the complexity of multiple compartments such as endosomes and nuclei, which themselves are dynamic and can undergo fusion and fission and exchange their content. Here, we show that stochastic pi calculus, a widely-used process algebra, is well suited for mapping surface and intracellular NP interactions and distributions. In stochastic pi calculus, each NP is represented as a process, which can adopt various states such as bound or aggregated, as well as be passed between processes representing location, as a function of predefined stochastic channels. We created a pi calculus model of gold NP uptake and intracellular movement and

compared the evolution of surface-bound, cytosolic, endosomal, and nuclear NP densities with electron microscopy data. We demonstrate that the computational approach can be extended to include specific molecular binding and potential interaction with signaling cascades as characteristic for NP-cell interactions in a wide range of applications such as nanotoxicity, viral infection, and drug delivery.

**Keywords** Nanoparticles · Nanotoxicity · Modeling · Process algebra · Delivery · Intracellular distribution

## Introduction

The growing use of nanotechnology in medicine and many novel products for daily use underscores the importance of studying the effect of nanoparticle (NP) uptake on human cells and organs. NPs have been used in a wide variety of biomedical applications (Stark 2011; Sperling et al. 2008; Murphy et al. 2008; Alivisatos et al. 2005), but simultaneously present an inherent risk to human health because of its ability to penetrate human tissue, circulate in the blood stream, and pass through cell membranes. Currently, there has been an increase in the number of studies dealing with the uptake and intracellular fate of NPs (Lévy et al. 2010; Nativo et al. 2008), which are mostly found enclosed within endo-lysosomal structures (Kneipp

---

M. P. D. Dobay (✉) · A. P. Alberola ·  
E. R. Mendoza · J. O. Rädler (✉)  
Faculty of Physics, Center for NanoScience, Ludwig-  
Maximilians University, Geschwister-Scholl-Platz 1,  
Munich, Germany  
e-mail: maria.pamela.david@physik.uni-muenchen.de

J. O. Rädler  
e-mail: joachim.raedler@physik.uni-muenchen.de

E. R. Mendoza  
Department of Computer Science, University of the  
Philippines, Diliman, Quezon City, Philippines

et al. 2006; Khan et al. 2007; Singh et al. 2010) and less frequently, as free particles in the cytosol (Nativo et al. 2008), mitochondria (Singh et al. 2010; Li et al. 2003; Xia et al. 2006), endoplasmic reticulum (Singh et al. 2010), or nucleus (Singh et al. 2010; Xia et al. 2006; Garcia-Garcia et al. 2005; Alberola and Radler 2009). Subcellular NP distribution data are good indicators of both therapeutic efficiency (Lévy et al. 2010; Sauer et al. 2009; Boeckle and Wagner 2006) as well as potential toxicity (Lévy et al. 2010; Sauer et al. 2009). There are also numerous reports on NP interference in signaling pathways, including apoptosis (Muller et al. 2010; Zhao et al. 2009; Park et al. 2008), inflammation (Muller et al. 2010), changes in gene expression, and protein folding (Khan et al. 2007). Most of the progress is in understanding the correlation between the composition of the adsorption layer of the NPs and uptake (Dell'Orco et al. 2010; Nel et al. 2009; Lundqvist et al. 2008; Cedervall et al. 2007; Lynch et al. 2007).

As early as in the 1990s, the need for *in silico* evaluation of adsorption, distribution, metabolism, excretion, and toxicity (ADMET) of new drugs, now naturally extended to NPs, was recognized (Haddish-Berhane et al. 2007; van de Waterbeemd and Gifford 2003). The current influx of data, however, is not always accompanied by a mechanistic understanding of the processes behind these, further bolstering the need for complementary quantitative and mechanistic models. A key requirement for such models is the possibility to cover broad spatio-temporal regimes that address all aspects of NP-cell interactions. Finally, the model should be able to handle stochastic behavior to account for the inherent variability between the NPs, as well as the role of the cell in response variability (Summers et al. 2011).

Currently, most NP-cell interaction models that address NP subcellular distributions are formulated using ordinary differential equations (ODEs) (Dell'Orco et al. 2010; Salvati et al. 2011; Dinh et al. 2007). ODE models, however, have known scalability limitations. Model extensions where variables with multiple interdependencies have to be changed, for instance, result in combinatorial explosion (Fisher and Henzinger 2007). This limitation has major implications for modeling multifunctional NPs, which are expected to cross-react extensively with the biological system. The inclusion of spatial information in ODE models would also require the definition of additional variables for each reactant in each of the specified compartments, is similarly cumbersome and non-intuitive.

Apart from ODE models, statistical and phenomenological models have also been used for investigating NP-cell interactions (Summers et al. 2011; Puzyn et al. 2011; Van Hoecke et al. 2009). Phenomenological models, while useful in describing how NP-cell interactions vary as a function of condition, do not yield information on why these variations occur. Furthermore, we have recently shown that the application of two different phenomenological models to the same data could result in radically different interpretations (Dobay et al. 2011).

The application of concepts from concurrency theory to biological modeling has opened a plethora of possibilities for creating mechanistic models of biological systems. Process algebra methods permit the description of communication, specific interaction, and synchronization between independent, mobile objects known as processes (Milner 1999; Priami 2009). Communication between processes is done through channels, with one process being the initiator of an action and another being the receiver. Channels have names, which allow communication to be specific to subsets of processes. Finally, it is possible to summarize reaction patterns as rules. These rules specify how reactants can be recognized, which is through channel names in  $\pi$ ; how the system is to be written as a result of a reaction. Owing to these characteristics, process algebra is particularly attractive for modeling biological systems.

In this article, we study the subcellular distribution of gold NPs by transmission electron microscopy (TEM), and use a stochastic process algebra model to simulate their entry and intracellular movement in living cells. Gold NPs have been chosen as a model system due to their stability, ease of preparation, and detection by TEM, as well as ready availability of experimental data from other groups (Nativo et al. 2008; Bartczak et al. 2011; Chithrani and Chan 2007; Chithrani et al. 2006). We develop a model scenario of NP interactions, which specifically includes membrane topology transitions via fusion and fission. Specifically, we use an implementation of the stochastic  $\pi$  calculus, SPiM (<http://research.microsoft.com/en-us/projects/spim/>), developed by Phillips and Cardelli (2004), at Microsoft Research, Cambridge which was developed for modeling biological processes. SPiM uses the Gillespie algorithm to choose the next reaction probability proportional to the base rate of reaction and reactant availability, as well as for

calculating its duration, insuring correct chemical kinetics (Gillespie 1977). We use both the NP-as-process and compartment-as-process abstractions, which enable us to obtain the subcellular distribution of NPs without the use of extra variables, as the case would have been in an ODE model. Model results demonstrate the potential of this technique for analyzing general NP-cell interactions, with a clear possibility of extension to more complex scenarios. We discuss the use of stochastic process algebra for general biological processes that involve particulate objects and cellular membrane compartments. The many parallels between NP uptake and other cellular uptake processes would ultimately make such models also useful in the areas of viral entry and gene delivery.

## Methods

### Cell culture

Cells were grown in Earl's MEM supplemented with 10 % FBS and 2 mM L-glutamine in a humidified atmosphere at 37 °C with 5 % CO<sub>2</sub> level. Cells were maintained at 85 % confluence, trypsinized, re-suspended in cell medium, and counted using a Neubauer counting chamber.

### Sample preparation for TEM

Human lung epithelial cells Beas-2B were seeded with a density of  $4 \times 10^4$  (Alivisatos et al. 2005) cells/well onto Millicell polycarbonate porous membranes. One day after seeding, 4-nm gold NPs 15 nM were added and incubated in serum for 6 h to a final concentration of 15 nM, following which cells were fixed overnight with Karnovsky's solution at 4 °C. Cells were washed with cacodylate buffer and post-fixed for 2 h with 1 % osmium tetroxide in cacodylate buffer at 4 °C. Ethanol was used to dehydrate the cells step by step before embedding in EPON resin (Poly/Bed812-BDMA Embedding Kit). Thin sections were stained using uranyl acetate and lead citrate. Cells are sectioned transversally with respect to the polycarbonate membrane on which they are attached. Gold NPs were kindly provided by Prof. Dr. W. Parak, AG Biophotonik, Marburg, Germany.

### Stochastic pi calculus model of NP entry and intracellular movement

We created an executable pi calculus model using the stochastic pi machine, v. 0.05 (SPiM, Phillips 2005, freely available from <http://research.microsoft.com/en-us/projects/spim/>). The model incorporates all reactions illustrated in Fig. 3a; rates and restrictions for each of the reactions are summarized in Fig. 3b. Each run corresponds to the dynamics of a cell that can take up a maximum of 2,000 NPs; results of 500 runs were processed by a Perl script. Run automation was facilitated by a simple shell script. The codes are available from <http://softmatter.physik.lmu.de/tiki-index.php?page=Downloads>.

## Results

### Subcellular distribution of gold NPs in epithelial cells as seen by TEM

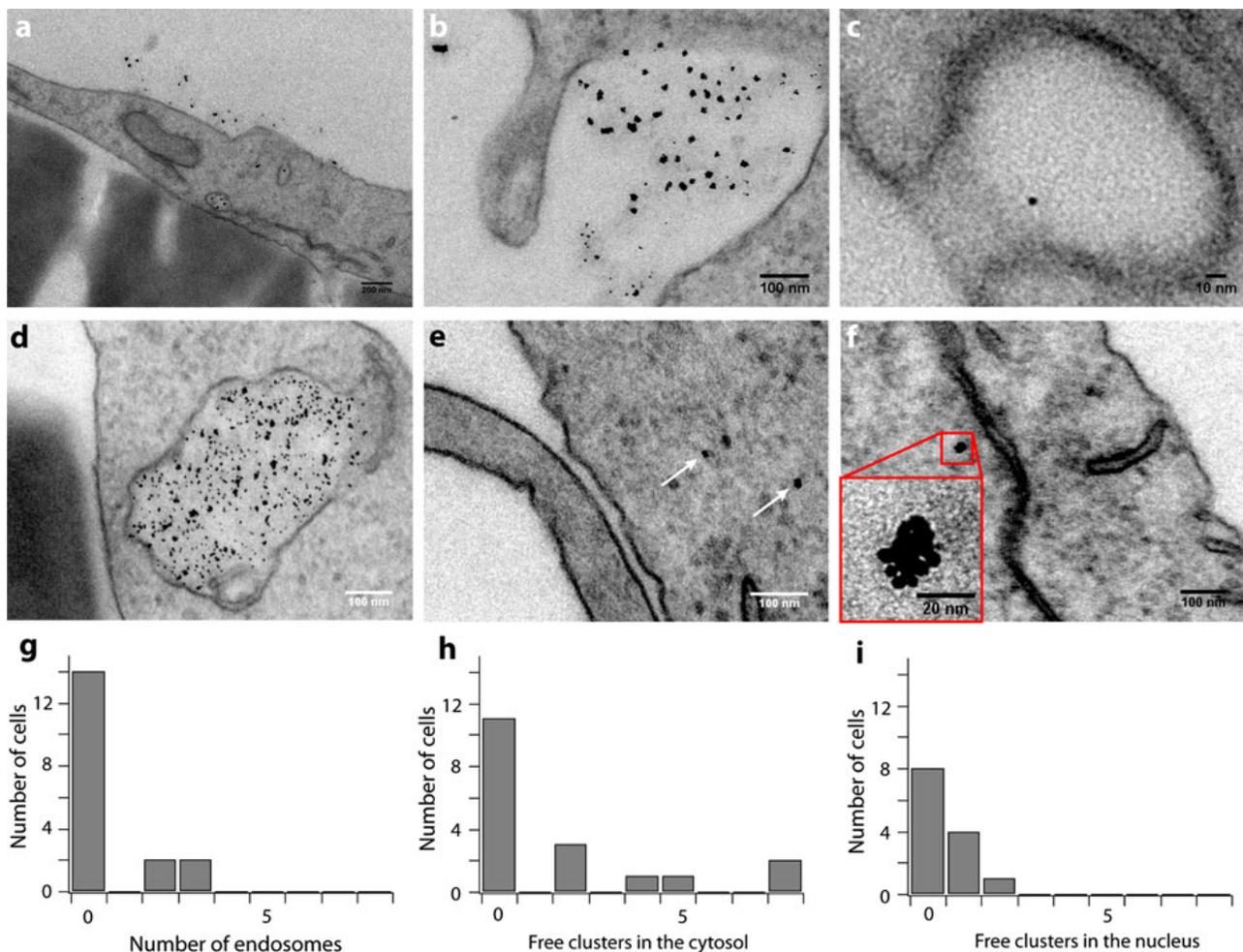
We performed in-house NP uptake experiments to demonstrate the uptake mechanisms and subcellular distribution of NPs. We studied lung epithelial cells incubated with 4-nm gold NPs imaged using TEM microscopy; these are visible in the cells as black contrast-rich, electron-dense spots. Typically, NPs aggregate after contact with cell medium (Fig. 1, micrographs a, b, d, e, and f). Consequently, NPs occur mostly as clusters inside cells: trapped in endosomes, free in the cytoplasm, and in the cell nucleus. Figure 1a shows an overview of a part of a cell where NPs and NP clusters are found close to the cell membrane and trapped in endosomes. We observe NPs entering cells via diverse endocytotic pathways, including macropinocytosis (Fig. 1b) and caveolae-mediated pathways (Fig. 1c). Note the difference in the magnification between micrographs 1b and 1c. While micropinocytosis leads to big endosomes that reach sizes up to 1 μm, caveolin-mediated endocytosis leads to smaller endosomes of around 60 nm (Luccardini et al. 2007). Since the different endocytotic pathways lead to endosomes with different sizes, the distribution of gold NPs in endosomes is non-uniform, with some endosomes containing a high number of NPs (Fig. 1d) than others, which can contain as little as a single NP (Fig. 1c). Figure 1e shows NP clusters, indicated by arrows, free in the cytoplasm. In the cell

nucleus, NP clusters with diameters ranging from 15–20 nm are found. One such cluster close to the nuclear membrane is shown in Fig. 1f. The histograms in the lower panel of Fig. 1 summarize the number of cells in terms of how much endosomes (Fig. 1g), free NP clusters in the cytosol (Fig. 1h), and NP clusters in the nucleus (Fig. 1i) they contain.

#### Implementation of NP interactions as a pi calculus model

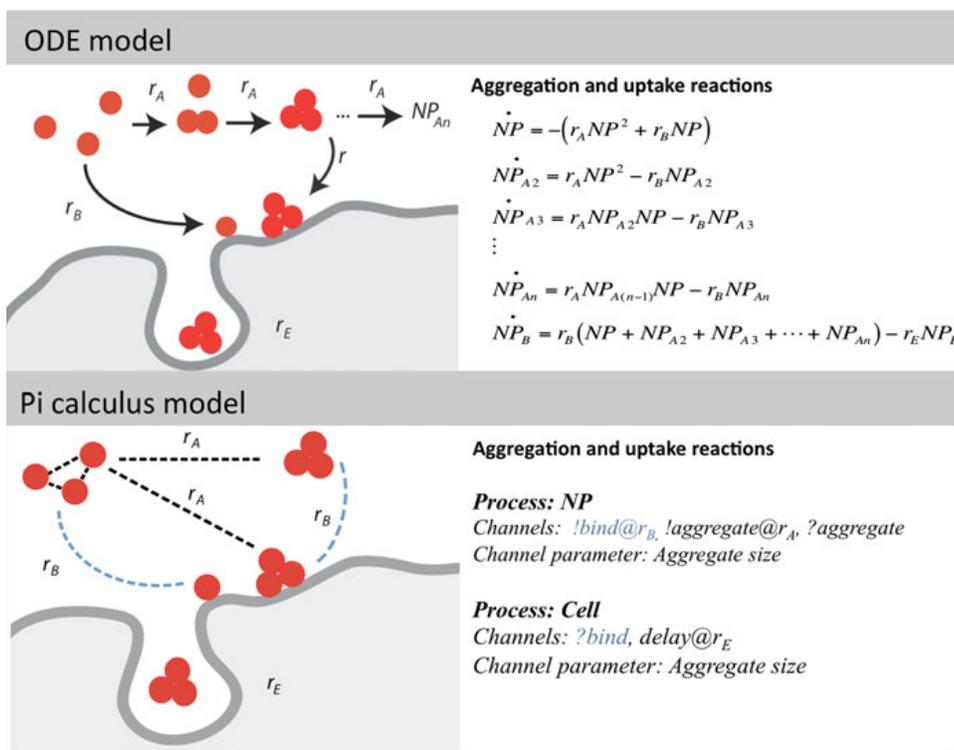
To understand this evolution of NP distribution inside cells, we created a simulation that details NP uptake

and intracellular movement. Our choice of pi calculus was motivated by scalability and expressivity and the necessity of including stochastic aspects of the system. To illustrate these points, we consider the first scenario of NP interactions, namely the aggregation of NP and their binding to the cell surface (see Fig. 2). An ODE model requires as many variables as there are NP species to distinguish between free NPs, NP clusters of size 2, 3, ...  $N$ , and their bound versions. This implementation is cumbersome, and involves an explosion of the number of required variables as clusters of larger size are included. Unlike in ODE models, where locations and clusters of NPs are necessarily treated as separate variables, in pi calculus, NPs are represented



**Fig. 1** TEM images of 4-nm gold NP uptake. A 15 nM solution of polymer-coated particles was incubated in with Beas-2B cells complete medium for 6 h before fixation and TEM imaging. **a** A typical overview of a part of a cell is shown. Cells are attached on a polycarbonate porous membrane and are cut transversally. NPs are found forming clusters outside and inside of the cell. NPs use different modes of entry into the cells,

including pinocytosis (**b**) and caveolae-mediated pathways (**c**). Inside cells, NPs are found localized in endosomes (**d**), free in the cytosol (**e**) and forming small clusters inside the cell nucleus (**f**). The number of endosomes (**g**), as well as the number of NP clusters in the cytosol (**h**) and the nucleus (**i**) observed in the examined sections are shown



**Fig. 2** Scalability problems in modeling NP aggregation and uptake using ODEs. An ODE model of NP aggregation and uptake requires the separate specification of binding and entry reactions for individual free (NP) and aggregated NPs of size 2, 3...  $N$  ( $NP_{A2}, NP_{A3}, \dots, NP_{An}$ ). Note the need for extra variables to track aggregates of different sizes. Apart from the obvious scalability problems that would arise if reactions involving the NP aggregates are introduced, one needs to have a priori knowledge of particle sizes for which variables have to be defined. In addition, a variable for NPs on the cell surface,  $NP_B$  has to be specified separately. If the size distribution of bound NPs is needed,  $n$  additional ODEs, each representing bound

aggregates of sizes  $1, \dots, n$ , have to be specified. In contrast, the equivalent pi calculus model only requires the definition of NPs as processes that could aggregate or bind to a cell. When an aggregate is formed, the size can be treated as an attribute of the process. Further scalability problems in ODE models will arise if the model is extended to include the intracellular location of a particle (Fig. 3a). Rate constants for aggregation ( $r_A$ ), binding ( $r_B$ ) are indicated in both models, and are assumed to be uniform for all reactions, regardless of aggregate size. Communication channels between processes (dashed lines) are similarly indicated for the pi calculus model, and are colored by channel name

as processes the state of which can change as a result of interactions with other processes. In any computational implementation, many of these processes are generated and concurrently updated in parallel (typically 2,000 processes in our simulations). NP processes have its state and size as attributes, and could be passed as parameters between processes representing location. Specifically, NPs are modeled as processes that can aggregate with each other through channels called “!aggregate” or “?aggregate” depending on whether it is an output (!) or input (?) channel, respectively. When an aggregation reaction occurs, the size attribute of the sender NP is passed through the !aggregate channel, and is added to the size attribute of the receiver NP. NP adhesion to the cell surface is modeled using the “binding” channel, and results in the concurrent

updates of the state of NP processes in an aggregate from “unbound” to “bound.” The corresponding code is shown in Fig. 2. A process calculus simulator such as SPiM executes the model by concurrently computing the states of multiple NPs using a stochastic algorithm. The program therefore keeps track of the fate of each NP at any time. This attribute permits both NP subcellular distribution and NP cluster size distributions to be easily obtained from the model. Process algebra is thus fundamentally distinct from ODEs in that the model of the system is not centered on concentrations, their relationships, and how they change over time. Instead, it is focused on the definition of processes that capture the behavior of the biological objects that they represent, and that can be directly used as objects in a simulation.

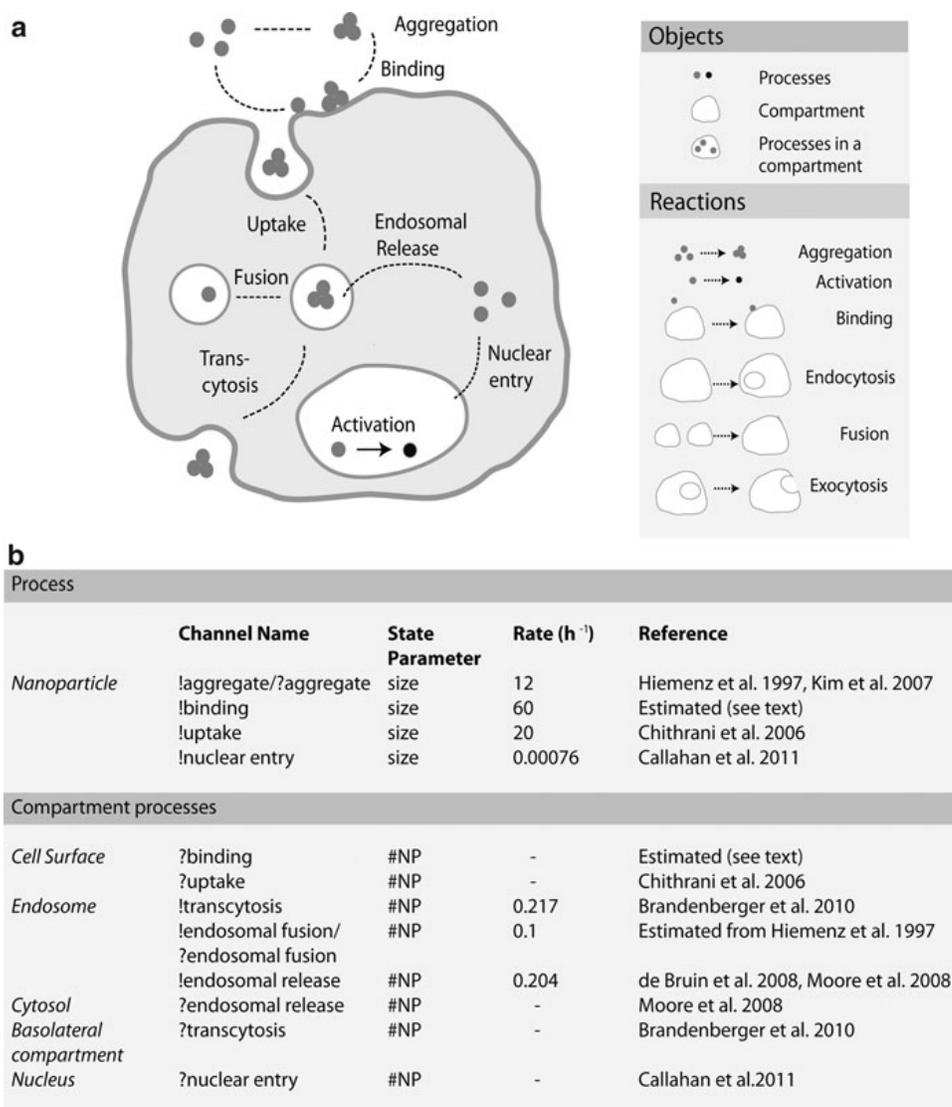
## Time evolution of NP uptake and subcellular distribution in silico

We proceeded to extend our model to simulate the entire gold NP cell uptake scenario. Including events downstream of binding requires the definition of the cellular locations as processes that can receive the NP process as a parameter. Various extensions of pi calculus which provide a more intuitive and elegant description of dynamic compartments have been developed. However, we developed the model using pi calculus constructs. These processes are involved in reactions that change the topology of cellular compartments and redistribute their contents accordingly; these include endocytosis, fusion, exocytosis, and dissolution-type reactions. Compartments may interact with regular processes or with other compartments. Figure 3a shows all the reactions that we considered in our model. Reactions can be between processes, as in NP aggregation; between a process and a compartment, like in binding; or between compartments, as in transcytosis. We assumed that both free and clustered NPs could be internalized after a two-phase interaction with the cell surface: the binding interaction results in a state change of the NPs involved, while an interaction on the ?uptake channel results in the creation of an endosome process. The number of internalized NPs per endosome is sent through the ?uptake channel. NPs could then escape from the endosome into the cytosol in a dissolution-type compartment reaction; NP complexes that escape the endosome are assumed to remain aggregated for simplicity. Alternately, the endosome could be transcytosed in an exocytosis-type reaction. Endosomes may also fuse with other endosomes, combining their NP contents. In all cases, an output on any of the endosome channels passes the number of NPs it contains onto the compartment with which it interacts. This implies that possible reactions per compartment are updated. All outputs on an endosome channel also terminate the specific instantiation of the endosome process, corresponding to the physical dissolution of the endosome that occurs in the cell. Of all the NPs, only those that are in the cytosol can enter the nucleus through the !nuclear entry channel; for clusters, a single output on this channel is sufficient to translocate all the NPs in it into the nucleus. Once in the nucleus, NPs can be eventually modeled to undergo an activation reaction, which is a generalization for diverse events that can range from

the release of NP cargo, or the initiation of an interaction of NPs with the genetic material. Although we do not include activation in our simulations, an elaborated form of this reaction could be used in evaluating either the efficiency of a NP designed to deliver cargo to the nucleus or to evaluate the cytotoxicity of a NP that can only be tolerated in the nucleus at small amounts.

Figure 3b provides a list of all processes and its corresponding channels corresponding to the model in Fig. 3a. Each channel in it is associated with rates, which, together with the number of processes bearing the channel, determine the probability that the channel event occurs. We endeavored to use realistic kinetic rates obtained from experiments using gold NPs reported in the literature, unless otherwise indicated (Fig. 3b). The aggregation rate was estimated as  $k_{\text{agg}} = k_r \cdot C/W$  based on the theory of coagulation in dilute dispersions (Hiemenz et al. 1997), where  $C$  is the molar concentration of NPs and  $W$  is the stability ratio.  $k_r$ , the rate constant of rapid coagulation is given by  $k_r = 4k_B T/3 \eta$ , where  $k_B$  is the Boltzmann constant,  $T$  is the temperature, and  $\eta$  is the viscosity of water. Assuming  $W$  to be  $5 \times 10$  (Alivisatos et al. 2005), as taken from (Kim et al. 2008), this yields the approximate aggregation rate of  $12 \text{ NP h}^{-1}$  at room temperature. For the binding rate, we made the crude assumption that binding is diffusion limited as well, and the rate chosen was five times higher than the aggregation reaction, taking the larger cell surface area into account. The uptake rate was taken to be  $20.0 \text{ NPs h}^{-1}$ , which was estimated from data for 50–74 nm gold particles (Chithrani et al. 2006). We address the issue of size-dependence of uptake rates more in detail at a later stage in this article. The transcytosis rate was obtained for 15-nm gold NPs, and was estimated to be of the order  $\sim 0.2\text{--}1.3 \text{ NPs h}^{-1}$ ; we took the lower limit of the rate. The endosome–endosome fusion rate was treated as an aggregation reaction, but with  $\eta$  replaced by local cytosol viscosity, which is estimated to be 20 times the viscosity of water (Leduc et al. 2011). This yields a value of around  $0.4 \text{ NP h}^{-1}$ , which is further reduced to  $0.1 \text{ NP h}^{-1}$  given that most of these events occur only around 40 min post-formation (Kjeken et al. 2004). For the rate of endosomal release, we took the lower limit of  $0.204 \text{ NP h}^{-1}$  reported for a polyethyleneimine (PEI) gene delivery vector, given that the NPs do not exhibit the proton sponge effect that aids in

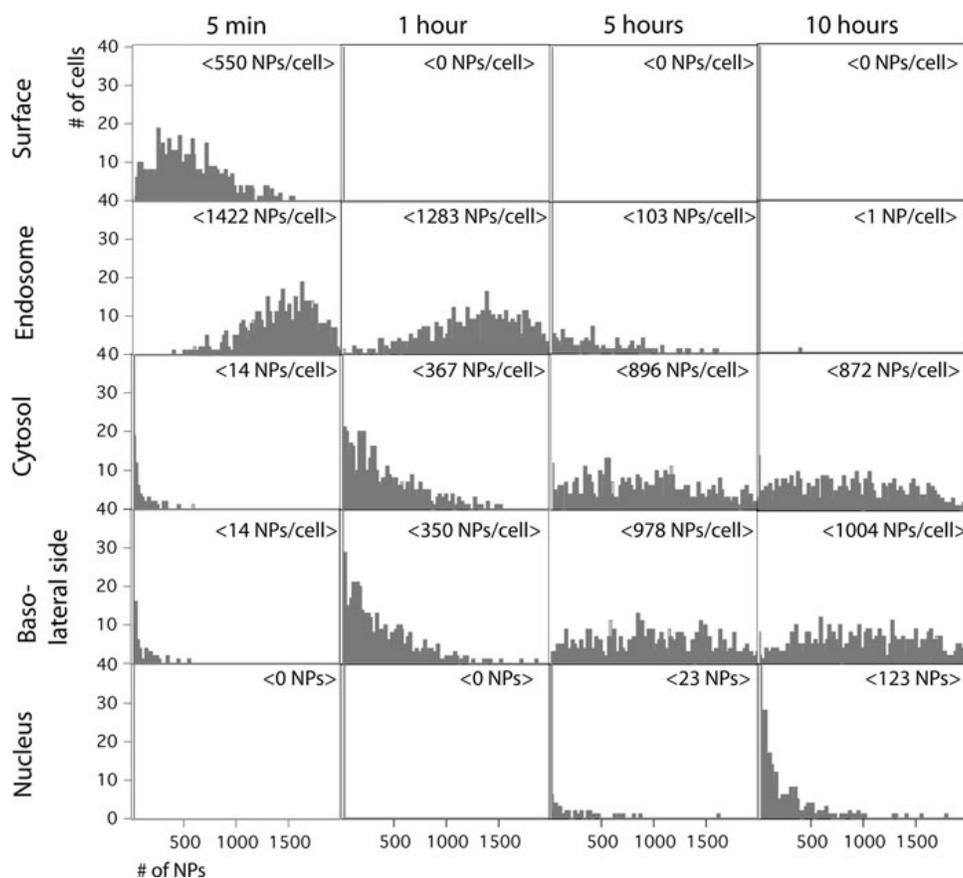
**Fig. 3** Schematic drawing of gold NP interaction and movement in cells. Single NPs can form aggregates of unrestricted size, or bind to a cell directly; NP aggregates can likewise bind the cell. On binding, both NPs and NP clusters are taken up by the cell into an endosome, which can then either fuse with other endosomes, release its content to the cytosol, or fuse with the cell membrane on the basolateral side of the cell. NPs can enter the nucleus through the entry of single NPs or NP aggregates in the cytosol (a). Both NPs and cell compartments are treated as processes in a pi calculus model of the system. NP processes are subsequently passed through channels between compartment processes. An overview of process channels and its parameters are shown in (b). Rates are taken from literature derived from gold NP, or equivalent experiments. All rates are associated with the output channel



PEI endosomal escape (de Bruin et al. 2008; Moore et al. 2008). Finally, the nuclear entry rate, both for NPs released was taken from HPMA copolymer particles having hydrodynamic diameters ranging from 2 to 4 nm (Callahan et al. 2009).

The model was run for 500 instances of cells, each of which started with maximum of 2,000 NPs in the extracellular milieu. The choice of 2,000 as the maximum number of NPs that can enter a cell was based on 14–74 nm gold NP experiments that placed this number to be at the order of ten (Murphy et al. 2008; Chithrani et al. 2006). Subcellular distributions obtained from this first model are shown in Fig. 4. The most prominent feature is the inward shift of the distribution over time. The first row of graphs shows how the distribution of NPs on the surface shifts to smaller values. At  $t = 30$  s, approximately 800 NPs

have already had some interaction with the cell. Of these, an average of 730 NPs are attached on the cell surface, while around 60 are already in endosomes (data not shown). After 5 min, most cells have an average of 550 NPs attached on the cell surface, while an average of 1,422 NPs are already in the endosome; these include both free and clustered NPs. Uptake is significantly higher than in Chithrani et al. (2006), where this level is expected to be obtained at around 1 h. After 5 h, all the NPs are predicted to be internalized. This is consistent with gold NP uptake results reported in literature, where uptake saturation for 14–74 nm particles begins at approximately 5.5 h (Chithrani et al. 2006). The second row shows how the NPs quickly fill endosomes. Endosomes then release their cargo either to the cytosol or to the basolateral site by exocytosis. Given that equal rates were assumed for



**Fig. 4** Simulation results of gold NP entry in 500 cells. The diagram shows the NP distributions over 500 cells at five different locations (surface, endosome, cytosol, baso-lateral side, and nucleus) as a function of time. At  $t = 0$ , 2,000 free NPs are assumed to be within the extracellular milieu of each of the 500 cells. At  $t = 5$  min, most of the NPs ( $\sim 1,500$  per cell) have been internalized, with around 25 % of the NPs remaining bound to the cell surface. At  $t = 1$  h, all the particles have been

these reactions, it is not unexpected that approximately half of the particles are lost to transcytosis after 5 h; only a very small percentage of the NPs are likewise predicted to be in the nucleus, consistent with experimental results (Fig. 1f, i). However, nuclear NPs are predicted to increase at 10 h, which is not expected from previous reports (Chithrani et al. 2006). Furthermore, our model predicts that the nuclear content is widely distributed, whereas experimental data indicate that nuclear clusters generally have diameters ranging 15–20 nm, roughly corresponding to complexes formed from between approximately 25–50 NPs from visual inspection (Fig. 1f). A likely reason for these discrepancies could be the fact that the initial uptake rates do not adequately discriminate between small and large clusters, and hence overestimates both the

taken up, 20 % of which have escaped into the cytosol. At  $t = 5$  h, cytosolic NPs reach a plateau of around 50 %. Only  $\sim 2$  % cells are predicted to contain particles in the nucleus. At  $t = 10$  h, the number of nuclear entry events increase significantly, and involve a wider size distribution of clusters than those observed in experiments. Note that a column sum of the averages for each location is equal to 2,000

number of cell and nuclear entry events, as well as the cluster size distribution.

#### Cluster size distribution and size-dependent rates

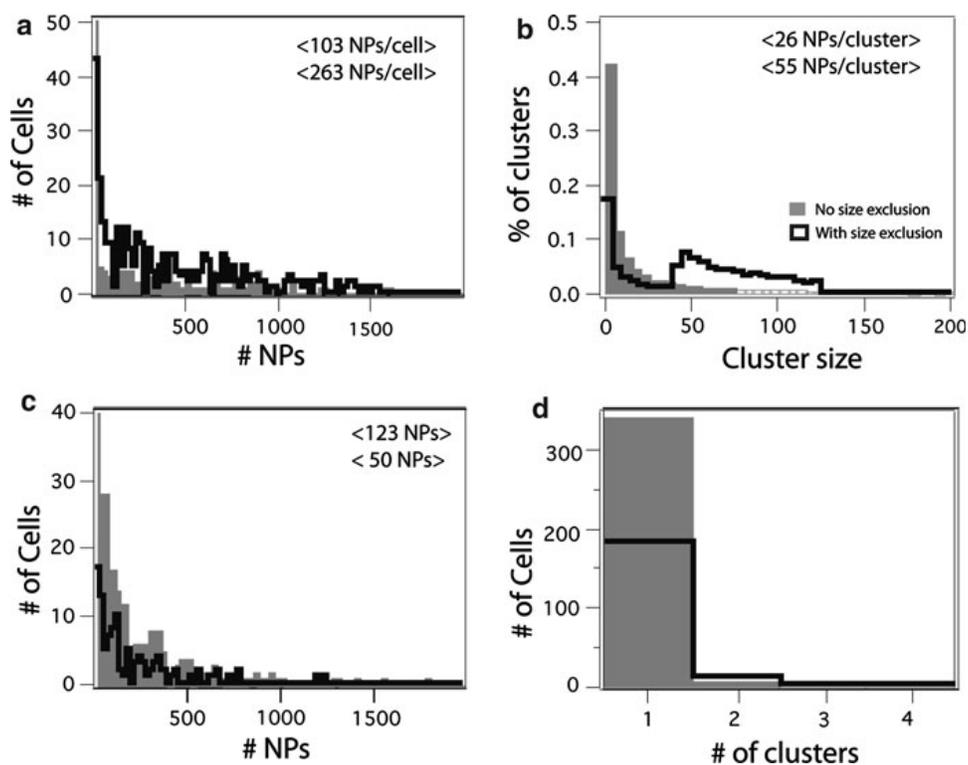
To account for size dependent uptake, rates can be changed based on size conditionals. As a first modification, we simulated a significantly higher uptake rate of  $622 \text{ NPs h}^{-1}$  for clusters of sizes 43–125 NPs, approximately corresponding to 15–20 nm clusters, which are known to be taken up optimally in comparison to smaller clusters, or to clusters exceeding a 100 nm diameter (Chithrani et al. 2006). Clusters comprised of 1–42 NPs, as well as those greater than 125 were assigned a lower uptake rate of  $1 \text{ NP h}^{-1}$ . In Fig. 5, we demonstrate the outcome of size-dependent

endosomal uptake. This refinement increases the number of NPs per endosome from an average of 103 NPs/cell to 263 NPs/cell (Fig. 5a). Evaluating the cluster size distributions of these endosomal NPs reveals that there is indeed a significant increase in larger NP clusters (Fig. 5b). The size-dependent enhanced endosomal uptake also causes a corresponding shift downstream in the number of nuclear NPs, which are now at an average of 50 NPs, down from the 123 NP average when there are no size constraints (Fig. 5c); this is now more consistent with experimental data. To check if the nuclear clusters are comprised of single clusters, or are comprised of more than one set of clusters, we checked the number of nuclear entry events. In both size-restricted and

unrestricted cases, most of the nuclear clusters result from single entry events that mostly involve NP clusters; size-restriction reduces the occurrence of nuclear entry from 70 to 40 %, which is in agreement with our data (Figs. 1f, 5d).

## Discussion

In this work, we introduced the concept of process algebra for modeling NP-cell interactions and highlighted the ease with which this approach allows mapping of distribution of NPs across different cellular compartments. At the same time, process algebra prevents the problem of combinatorial



**Fig. 5** Effect of size restrictions on uptake and nuclear localization of NPs. Based on sphere packing calculations and image analysis results, NP aggregates of size 43–125 were assigned an uptake rate of  $622 \text{ NPs h}^{-1}$ , while other aggregates were assigned a suboptimal rate of  $1 \text{ NP h}^{-1}$ , based on previous experimental results. Incorporating the effect of aggregate size on uptake rate, results in a significant reduction of single particles and small clusters taken up in endosomes. These are reflected in both the number of NPs in endosomes, which increases from an average of 251 NPs/cell to 263 NPs/cell (a), as well as the average cluster size, which increases from 26 to 55 NPs/cluster (b). The effects of cluster size restrictions on NP

uptake are also reflected in the sizes of nuclear-localized clusters, which goes down from 123 to 50 NPs (c), a value more consistent with experimental results, where clusters range from 15 to 20 nm. Size restriction, however, does not generally affect the distribution shape of nuclear entry events, which indicate that most nuclear NPs occur from single entry events of NP clusters (d). Instead, restriction decreases the total number of nuclear entry events from 70 to around 40 %, which is also more consistent with TEM results. Endosome distributions were obtained at  $t = 5 \text{ h}$ , while nuclear distributions were obtained at  $t = 10 \text{ h}$ . The total number of aggregates was for 500 cells that can internalize a maximum of 2,000 NPs each

explosion, which occurs when NP aggregation is considered. In particular, we described a stochastic pi calculus model that simulates the subcellular distribution of NPs as a function of its binding affinity, size, (Milner 1999) and aggregation rate, as well as of the topological changes of the compartments where it can localize. Recent developments in NP research and practical application has led to burgeoning amounts of data, including subcellular distribution information, but an understanding of the mechanisms governing NP behavior does not appear as forthcoming. One of the important first steps in this direction is the creation of a computational model from which one could glean the plausibility and/or adequacy of assumptions regarding the system (Kitano 2002).

A critical step in creating any computational model is parameter estimation; these parameters include both reaction rates and reactant stoichiometry. When possible, initial parameters are obtained from experimental data. In this case, one could test the model structure better, and deviations between the model fit and the data could be attributed to the proper source of error (Faller et al. 2003). It is rarely the case, however, that all required parameters are available. Assumptions often have to be made about the order of magnitude of the reaction rates, although it is possible to calibrate the model to fit the experimental results in a process called parameter estimation (Moles et al. 2003). Once parameters are estimated from the data, these are statistically validated against all observed dynamic behaviors of the system being modeled (Faller et al. 2003).

For our model, we used rates for gold NPs, when available, as well as from other particle as initial values for our simulations. The first subcellular distributions that we obtained from the model, however, were already generally consistent with experimental data. A difference that we decided to address is the result for nuclear NP cluster sizes. Given that it is relatively straightforward to access parameters in the model, we incorporated the non-linear dependence of uptake rate on particle diameter. This modification results in a significant reduction in the uptake of single NPs and small clusters (Fig. 5), as well as a shift of the average nuclear content size from 8 to 32 NPs, which are now more consistent with data (Fig. 5c).

An important point that we made by comparing our model with experimental results is that a pi calculus model is very useful for testing hypotheses. Changes

are both easily incorporated and simulated; the states of a process can also be accessed easily. We exploited the latter feature to test the effect of having size-dependent restrictions on uptake rates. This small change in the model significantly improved the consistency between simulation results and experimental data. A remaining discrepancy, however, is the breadth size distribution of NPs in the nucleus. Again, this is possibly due to the absence of size restrictions on the clusters that can enter the nucleus. If this is the case, then a similar size exclusion condition based on the nuclear pore diameter should be sufficient to handle the discrepancy.

Models can be further extended by adding channels for new interactions to existing processes, or defining a new process with which old processes can interact with. These do not often require extensive changes to existing code. It must be borne in mind that the current model is still comparatively simple, without details on NP-surface properties, or its possible interactions with proteins, most notably, corona formation (Cedervall et al. 2007). Physicochemical properties, which can be related to cytotoxic activity by the quantitative structure-activity relationship (QSAR) method, are not specified, either (Puzyn et al. 2011). Given that these factors are very important in evaluating NP cytotoxicity, it is particularly important that there is an efficient way by which these could be incorporated into a computational model of NP-cell and NP-protein interaction. Furthermore, it is of interest in the medical field to extend NP modeling to predictive simulations of intracellular distribution of drug or gene-loaded NPs. Gene delivery is a particularly appropriate example for quantitative NP modeling since the activity of successfully delivered NPs can be easily followed by expression of GFP fusion proteins (Schwake et al. 2010); this process could be included as the activation endpoint indicated in Fig. 3.

For any of these extensions, stochastic pi calculus models present a significant potential, as it allows large systems to be modeled incrementally. The specific implementation that was used to implement the model, SPiM v.0.05, requires code written in SPiM language, the syntax of which is slightly different from pi calculus. A web-based, platform-independent application prototype with a full graphical user interface ([http://lepton.research.microsoft.com/VisualSPiM/frame\\_silverlight.html](http://lepton.research.microsoft.com/VisualSPiM/frame_silverlight.html)) is also available, but run automation in this version is still under development.

Similarly, a drag-and-drop interface for code generation is being developed in the group of Phillips to preclude the need for users to learn the SPiM language to create a pi calculus model (Service 2011). These additional improvements to SPiM would make it even easier to evaluate the potential of pi calculus models in NP-related applications. As indicated previously, there are other process algebras that were specifically designed for handling compartments, such as Brane calculus (Cardelli 2005) and its variants (Danos and Pradalier 2005; David et al. 2009); BioAmbients (Schwake et al. 2010; Regev et al. 2004) and Beta-binders (Priami and Quaglia 2005). Currently, scalable and automatable tools for these are under development. These would eventually provide a more elegant and intuitive way of writing more complex, compartment-based models.

## Conclusion

We demonstrated the potential of pi calculus for modeling problems involving reactant state changes and topological transformations, which are expected to be key features in NP cytotoxicity modeling, and which are likewise expected to be the source of combinatorial explosion problems in equivalent ODE models. The current gold NP model that we presented sufficiently captures intracellular distribution patterns from experiments. In the future, this can be easily extended to detail how its distribution and properties affect how it interacts with intracellular components.

**Acknowledgments** This work was initially funded by the EU-FP6 project NanoInteract (contract 033231) and was vitally inspired by the EU-FP7 project “Transkinetics”. Further support was received from the excellence cluster nanosystems initiative Munich (NIM) and the Center for NanoScience (CeNS). MPD gratefully acknowledges the Deutscher Akademischer Austausch Dienst for her Ph. D scholarship.

## References

- Alberola AP, Radler JO (2009) The defined presentation of nanoparticles to cells and their surface controlled uptake. *Biomaterials* 30:3766–3770
- Alvisatos AP, Gu W, Larabell C (2005) Quantum dots as cellular probes. *Ann Rev Biomed Eng* 7:55–76
- Bartczak D, Sanchez-Elsner T, Louafi F, Millar TM, Kanaras AG (2011) Receptor-mediated interactions between colloidal gold nanoparticles and human umbilical vein endothelial cells. *Small* 7:388–394

- Boeckle S, Wagner E (2006) Optimizing targeted gene delivery: chemical modification of viral vectors and synthesis of artificial virus vector systems. *AAPS J* 8:E731–E742
- Callahan J, Kopeckova P, Kopecek J (2009) Intracellular trafficking and subcellular distribution of a large array of HPMA copolymers. *Biomacromolecules* 10:1704–1714
- Cardelli L (2005) Brane calculi—interactions of biological membranes. *Lect Notes Comput Sci* 3082:257–278
- Cedervall T et al (2007) Understanding the nanoparticle-protein corona using methods to quantify exchange rates and affinities of proteins for nanoparticles. *Proc Natl Acad Sci USA* 104:2050–2055
- 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
- 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
- Danos V, Pradalier S (2005) Projective Brane calculus. *Lect Notes Comput Sci* 3082:134–148
- David MPC, Bantang JY, Mendoza ER (2009) A projective Brane calculus with activate, bud and mate as primitive actions. *Trans Comput Syst Biol XI* 5750:164–186
- de Bruin KG et al (2008) Dynamics of photoinduced endosomal release of polyplexes. *J Control Release* 130:175–182
- Dell’Orco D, Lundqvist M, Oslakovic C, Cedervall T, Linse S (2010) Modeling the time evolution of the nanoparticle-protein corona in a body fluid. *PLoS One* 5(6):e10949
- 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
- Dobay MP, Dobay A, Bantang J, Mendoza E (2011) How many trimers? Modeling influenza virus fusion yields a minimum aggregate size of six trimers, three of which are fusogenic. *Mol Biosys* 7(10):2741–2749
- Faller D, Klingmuller U, Timmer J (2003) Simulation methods for optimal experimental design in systems biology. *Simul Trans Soc Model Simul* 79:717–725
- Fisher J, Henzinger TA (2007) Executable cell biology. *Nat Biotechnol* 25:1239–1249
- Garcia-Garcia E et al (2005) A methodology to study intracellular distribution of nanoparticles in brain endothelial cells. *Int J Pharm* 298:310–314
- Gillespie DT (1977) Exact stochastic simulation of coupled chemical-reactions. *J Phys Chem* 81:2340–2361
- Haddish-Berhane N, Rickus JL, Haghghi K (2007) The role of multiscale computational approaches for rational design of conventional and nanoparticle oral drug delivery systems. *Int J Nanomed* 2:315–331
- Hiemenz PC, Rajagopalan R (eds) (1997) Principles of colloid and surface chemistry, 3rd edn. Marcel Dekker, New York
- Khan JA, Pillai B, Das TK, Singh Y, Maiti S (2007) Molecular effects of uptake of gold nanoparticles in HeLa cells. *ChemBioChem* 8:1237–1240
- Kim T, Lee CH, Joo SW, Lee K (2008) Kinetics of gold nanoparticle aggregation: experiments and modeling. *J Colloid Interface Sci* 318:238–243
- Kitano H (2002) Systems biology: a brief overview. *Science* 295:1662–1664

- Kjeken R et al (2004) Fusion between phagosomes, early and late endosomes: a role for actin in fusion between late, but not early endocytic organelles. *Mol Biol Cell* 15:345–358
- Kneipp J, Kneipp H, McLaughlin M, Brown D, Kneipp K (2006) In vivo molecular probing of cellular compartments with gold nanoparticles and nanoaggregates. *Nano Lett* 6:2225–2231
- Leduc C, Jung JM, Carney RR, Lounis B, Stellacci F (2011) Direct investigation of intracellular presence of gold nanoparticles via photothermal heterodyne imaging. *ACS Nano* 5:2587–2592
- Lévy R, Shaheen U, Cesbron Y, Sée V (2010) Gold nanoparticles delivery in mammalian live cells: a critical review. *Nano Rev* 1
- Li N et al (2003) Ultrafine particulate pollutants induce oxidative stress and mitochondrial damage. *Environ Health Perspect* 111:455–460
- Luccardini C et al (2007) Getting across the plasma membrane and beyond: intracellular uses of colloidal semiconductor nanocrystals. *J Biomed Biotechnol* 2007:68963
- Lundqvist M et al (2008) Nanoparticle size and surface properties determine the protein corona with possible implications for biological impacts. *Proc Natl Acad Sci USA* 105:14265–14270
- Lynch I et al (2007) The nanoparticle—protein complex as a biological entity; a complex fluids and surface science challenge for the 21st century. *Adv Colloid Interface* 134–35:167–174
- Milner R (1999) *Communicating and mobile systems: the pi-calculus*. Cambridge University Press, Cambridge, England
- Moles CG, Mendes P, Banga JR (2003) Parameter estimation in biochemical pathways: a comparison of global optimization methods. *Genome Res* 13:2467–2474
- Moore NM, Sheppard CL, Barbour TR, Sakiyama-Elbert SE (2008) The effect of endosomal escape peptides on in vitro gene delivery of polyethylene glycol-based vehicles. *J Gene Med* 10:1134–1149
- Muller L et al (2010) Oxidative stress and inflammation response after nanoparticle exposure: differences between human lung cell monocultures and an advanced three-dimensional model of the human epithelial airways. *J R Soc Interface* 7(Suppl 1):S27–S40
- Murphy CJ et al (2008) Gold nanoparticles in biology: beyond toxicity to cellular imaging. *Acc Chem Res* 41:1721–1730
- Nativo P, Prior IA, Brust M (2008) Uptake and intracellular fate of surface-modified gold nanoparticles. *ACS Nano* 2:1639–1644
- Nel AE et al (2009) Understanding biophysicochemical interactions at the nano-bio interface. *Nat Mater* 8:543–557
- Park EJ et al (2008) Oxidative stress and apoptosis induced by titanium dioxide nanoparticles in cultured BEAS-2B cells. *Toxicol Lett* 180:222–229
- Phillips A, Cardelli L (2004) A correct abstract machine for the stochastic pi-calculus. In *Proceedings of BioCONCUR 2004*. London
- Priami C (2009) Algorithmic systems biology. *Commun ACM* 52:80–88
- Priami C, Quaglia P (2005) Beta binders for biological interactions. *Lect Notes Comput Sci* 3082:20–33
- Puzyn T et al (2011) Using nano-QSAR to predict the cytotoxicity of metal oxide nanoparticles. *Nat Nanotechnol* 6:175–178
- Regev A, Panina EM, Silverman W, Cardelli L, Shapiro E (2004) BioAmbients: an abstraction for biological compartments. *Theor Comput Sci* 325:141–167
- Salvati A et al (2011) Experimental and theoretical comparison of intracellular import of polymeric nanoparticles and small molecules: toward models of uptake kinetics. *Nanomedicine* 7(6):816–826
- Sauer AM et al (2009) Dynamics of magnetic lipoplexes studied by single particle tracking in living cells. *J Control Release* 137:136–145
- Schwake G et al (2010) Predictive modeling of non-viral gene transfer. *Biotechnol Bioeng* 105:805–813
- Service RF (2011) Computer models coming soon to a lab near you: drag-and-drop virtual worlds. *Science* 331:669–671
- Singh S, Kumar A, Karakoti A, Seal S, Self WT (2010) Unveiling the mechanism of uptake and sub-cellular distribution of cerium oxide nanoparticles. *Mol BioSyst* 6:1813–1820
- Sperling RA, Rivera Gil P, Zhang F, Zanella M, Parak WJ (2008) Biological applications of gold nanoparticles. *Chem Soc Rev* 37:1896–1908
- Stark WJ (2011) Nanoparticles in biological systems. *Angew Chem Int Ed* 50:1242–1258
- Summers HD et al (2011) Statistical analysis of nanoparticle dosing in a dynamic cellular system. *Nat Nanotechnol* 6:170–174
- van de Waterbeemd H, Gifford E (2003) ADMET in silico modelling: towards prediction paradise? *Nat Rev Drug Discov* 2:192–204
- Van Hoecke K et al (2009) Fate and effects of CeO<sub>2</sub> nanoparticles in aquatic ecotoxicity tests. *Environ Sci Technol* 43:4537–4546
- Xia T et al (2006) Comparison of the abilities of ambient and manufactured nanoparticles to induce cellular toxicity according to an oxidative stress paradigm. *Nano Lett* 6:1794–1807
- Zhao J et al (2009) Titanium dioxide (TiO<sub>2</sub>) nanoparticles induce JB6 cell apoptosis through activation of the caspase-8/bid and mitochondrial pathways. *J Toxicol Environ Health A* 72:1141–1149

# Appendix D

## Original publication P4

Andreas Hohner, Maria Pamela David, Joachim O. Rädler

**Controlled solvent-exchange deposition of phospholipid membranes onto solid surfaces**

*Biointerphases* **5**: 1 - 8, 2010.



# Controlled solvent-exchange deposition of phospholipid membranes onto solid surfaces

Andreas O. Hohner, Maria Pamela C. David, and Joachim O. Rädler<sup>a)</sup>

*Centre for Nanoscience and Fakultät für Physik, Ludwig-Maximilians-Universität, Geschwister-Scholl-Platz 1, München D-80539, Germany*

(Received 30 November 2009; accepted 11 January 2010; published 3 March 2010)

Controlled deposition of lipid bilayers plays a key role in creating supported membranes for biosensing devices and biophysical cell studies. The authors adopt a solvent-exchange method in order to deposit a phospholipid bilayer on solid substrates. The basic concept of deposition is to dissolve phospholipids in isopropanol-water mixtures and to increase water content gradually. Shortly before the onset of the micelle-to-vesicle transition, a lipid bilayer nucleates at the solid surface. They investigate the bulk phase behavior and surface coverage using small angle x-ray scattering and attenuated total reflection-Fourier transform infrared spectroscopy. They find a sequence of transitions from inverted-monomeric-micellar and vesicle phases correlating with an increasing amount of lipid on the adsorption layer. Supported lipid bilayers, prepared using this approach, are homogeneous and fluid. © 2010 American Vacuum Society.  
[DOI: 10.1116/1.3319326]

## I. INTRODUCTION

Many of the structural and dynamical properties of lipid bilayers, which characterize the unique properties of biological membranes, can be preserved when deposited on solid surfaces. As advocated in a seminal article by Sackmann,<sup>1</sup> the scientific and practical applications of supported membranes are versatile, and include applications such as biophysical model systems, biosensors or phantom cells. In general, molecules and their molecular interactions are more easily detected or imaged when anchored to surfaces. Supported membranes can be prepared as lipid monolayer or bilayers. They can be chemically grafted to the solid or adsorbed due to unspecific surface interactions. These have the potential to be used in various ways in combination with structured surfaces.<sup>2–8</sup> For instance, it is possible to confine fluid corals to chemically defined fields.<sup>6</sup> This allows the creation of integrated devices, such as lipid arrays or parallel assays in lipid-based chips for analytical and diagnostic applications.<sup>2,4,5</sup> Moreover, supported membranes can be combined with electrodes for monitoring membrane impedance of pores,<sup>7,9</sup> or with semiconductor technology for the application of lateral fields for membrane based electrophoresis.<sup>2,10,11</sup> Since supported membranes remain fluid, adsorbed macromolecules can easily be rearranged, permitting DNA molecules to be prepared in a stretched state.<sup>12</sup> Setups required for observing and manipulating the spatial organization of cell-model membrane interactions benefit from supported membranes in that the interacting components are well presented, such as those involved in the specific immunological recognition processes of T-cells.<sup>13</sup> An important role of lipids is also expected in future nanobiosystems. Here, supported membranes can be designed with nanoscale precision as nanofluidic lanes for transporting and monitoring single molecules.<sup>8,14</sup>

There are various methods by which supported lipid bilayers (SLBs) can be deposited on surfaces. SLBs are commonly prepared by vesicle fusion, a method developed in the McConnell laboratory,<sup>15,16</sup> where sonicated vesicles in contact with an appropriately cleaned glass surface rupture and spread out to form a continuous bilayer.<sup>17</sup> The difference in surface free energy between a bare and lipid bilayer-coated glass surfaces under water is the driving force that drags the bilayer from an area of excess lipid to uncovered areas.<sup>18,19</sup> Other methods to prepare supported membranes include the Langmuir–Blodgett deposition,<sup>15</sup> spin coating,<sup>20</sup> evaporation induced self-assembly,<sup>21</sup> and more recently, dip-pen technology.<sup>8</sup> Each of these methods has inherent advantages with respect to the applications in which they will be used. For instance, the Langmuir–Blodgett deposition allows the deposition of asymmetric bilayers,<sup>22</sup> dip-pen technology facilitates the fabrication of biomolecular arrays,<sup>8</sup> spin coating enables the quick generation of highly oriented, homogeneous bilayer stacks with defined thickness,<sup>20</sup> and evaporation induced self-assembly permits the creation of patterned nanocomposites of dissimilar materials.<sup>21</sup> In some applications, exposure of the sample to air should be avoided, and solution-based deposition of supported bilayers is preferred;<sup>22</sup> such procedures would also be suitable for a wider class of surfaces, including hydrophobic surfaces<sup>23</sup> or, in principle, also solid particles such as silica or glass beads.<sup>24,25</sup> Tiberg *et al.*,<sup>26</sup> for example, reported a simple method of this kind for preparing model lipid bilayers by coadsorption with a nonionic surfactant.

In this article, we present a solvent exchange method to prepare supported membranes. In this process, the quantity of the organic solvent, in which the lipids are dissolved, is continuously varied by addition of water. The approach is inspired by the reverse-phase evaporation method, where liposomes are produced through the slow removal of organic solvent from a water-solvent mixture.<sup>27</sup> Specifically, we in-

<sup>a)</sup>Electronic mail: joachim.raedler@physik.uni-muenchen.de

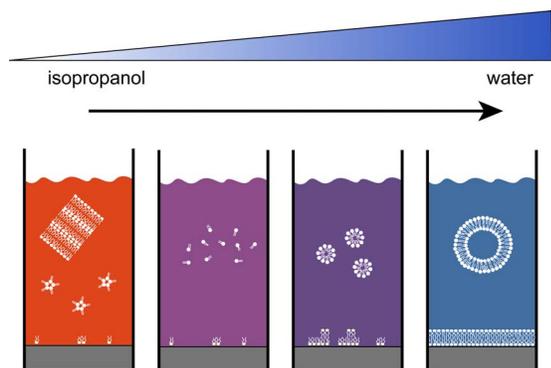


FIG. 1. (Color online) Principle of solvent exchange-induced physisorption. Lipids dissolved in pure alcohol are not adsorbed onto a  $\text{SiO}_2$  substrate. As water content increases, micelles begin to form, which can also adhere to the substrate. In pure water, lipid bilayers exist both on the substrate, as well as in solution (as liposomes).

investigate the lyotropic phases of lipids in isopropanol-water mixtures and the degree of lipid adsorption to solid surfaces as a function of increasing water content (i.e., solvent exchange), as schematically depicted in Fig. 1. Isopropanol was chosen from the commonly used solvents because of its complete miscibility with water. The lipids 1,2-dimyristoyl-*sn*-glycero-3-phosphocholin (DMPC) and 1,2-dimyristoyl-*sn*-glycero-3-trimethylammoniumpropan (DMTAP) were chosen to investigate the effect of lipid charges on structure formation on negatively charged surfaces. We present data obtained using small-angle x-ray scattering (SAXS) and attenuated total reflection (ATR) infrared spectroscopy to determine the bulk phase behavior and the degree of surface adsorption. A micelle-to-bilayer transition was found to occur at water contents between 80% and 90%. It is shown that a continuous fluid lipid bilayer deposits at glass surfaces when the water content of a lipid/alcohol/water mixture is gradually increased across this transition point.

## II. MATERIALS AND METHODS

### A. Reagents

DMPC and a cationic lipid, DMTAP, as well as its equivalents with deuterated fatty acid chains (d54-DMPC and d54-DMTAP), were purchased from Avanti Polar Lipids (Alabaster, USA). A lipid with a longer chain, 1-stearoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (SOPC, 18:0/18:1—phosphatidyl choline), was also obtained from Avanti Polar Lipids. All lipids were delivered dissolved in chloroform. Isopropanol was obtained from Fluka (Buchs, Switzerland). All compounds were used without further purification.

### B. Lipid preparation

Lipid samples were prepared by evaporation of the chloroform under a nitrogen flow and subsequent storage of the dried lipid in vacuum overnight. Degassed water and isopropanol were mixed by their volume properties and added to

the lipid. For Fourier transform infrared (FTIR) measurements, the lipids were dissolved in a mixture with the final water:isopropanol volume ratio. Samples for SAXS measurements were prepared similarly in cylindrical tubes with an outer diameter of 1.0 mm and wall thickness of 0.01 mm (Hilgenberg, Malsfeld, Germany).

For the continuous bleaching experiments, 2 mg lipids were tagged with 0.25 mol % Oregon green (Invitrogen) and dried overnight, as previously described.<sup>24</sup> Lipids were then initially dissolved in 1 ml of 50:50 (volume ratio) water:isopropanol solution and introduced into 100  $\mu\text{l}$  Ibidi bottomless chambers (Ibidi, Munich) fitted with a silicon oxide surface (Roth, Germany). These surfaces were prepared by sonication in 2% Hellmanex for 15 min, followed by washing with sonication in distilled, de-ionized water (Mill-Q, Millipore Co., USA). All surfaces were dried under a nitrogen stream prior to use. Samples were slowly titrated with water, with 10 min intervals prior to the addition of water that changes the content 5% at a time, until a 90% content is reached. After the last titration, this system was allowed to equilibrate for 30 min prior to washing and measurement. Samples were washed at least five times prior to measurement to ensure the removal of remaining vesicles.

### C. FTIR

ATR-FTIR experiments were performed with a Nicolet 60SXR FTIR spectrometer with a self-developed horizontal ATR flow cell. The cell consists of a polycarbonate window to avoid the formation of air bubbles and an aluminum crystal holder for a water-based temperature control.

To imitate a glasslike substrate a silicon crystal with a native  $\text{SiO}_2$  surface was used as the ATR crystal. The silicon crystal's surface was purified by storing it in a 2% sodium dodecyl sulfate solution for one day and cleaning it in a mixture of 70% nitric acid, 10% hydrogen peroxide, and 20% water for 1 h. The crystal was rinsed under water and dried for one day. Infrared spectra  $I_B(\nu)$  of the dry cell,  $I_R(\nu)$  of the water/isopropanol filled cell, and  $I_S(\nu)$  of the cell filled with lipid and solvent were measured. The corrected spectra were calculated by

$$I = \frac{I_S}{I_B} - \kappa \frac{I_R}{I_B}.$$

The coefficient  $\kappa$  was adjusted to minimize the water bands in the spectra.<sup>28</sup> For each lipid mixture, a part of the sample and the reference measurements are used in order to prevent the retention of any background after the correction. Lipids used had a concentration of 1 mg/ml. A 15 min equilibration period is allotted after samples are placed in the measurement chamber; this also allows the water vapor and  $\text{CO}_2$  in the chamber to be reduced.<sup>29</sup> To determine the absorption intensities, the spectra were matched against a sum of six Gaussian curves, with their position and width determined from reference measurements involving higher lipid concentrations. For the analysis, a prefactor was varied and data were fitted with a Gaussian-type baseline correction.

## D. SAXS experiments

SAXS experiments were carried out at the high brilliance beamline ID2 at the ESRF (Grenoble, France). X-ray scattering intensity at 0.15 nm was measured by a multiwire proportional gas-filled detector and radial integrated. The investigated  $q$  range for SAXS was  $0.2\text{--}5\text{ nm}^{-1}$ . The sample environment allowed temperature control within  $0.1\text{ }^\circ\text{C}$  through a Peltier water bath setup. The rotational symmetric scattering data were integrated radially and fitted with a model of three definite electron densities for solvent, acyl chains, and headgroup. Scattering curves were fitted in the wave vector range between 0.8 and  $5\text{ nm}^{-1}$  without background correction. The headgroup of the lipids was assumed to consist of an electron density  $\rho_H = 0.45\text{ e}^-/\text{\AA}^3$  and acyl chains of  $\rho_c = 0.17\text{ e}^-/\text{\AA}^3$ .<sup>30,31</sup> The electron densities of isopropanol-water mixtures range from  $\rho_0 = 0.157\text{ e}^-/\text{\AA}^3$  (0% water) to  $0.266\text{ e}^-/\text{\AA}^3$  (100% water).

## E. Determination of the lyotropic phases

The multilamellar phases were characterized by the position of the Bragg peaks  $d = 2\pi/q$  with no further analysis of the structure and form factor.<sup>32–34</sup> Isotropic phases were analyzed by fitting the intensity  $I(q) = F(q)^2$  with the theoretical form factors  $F(q)$ , corresponding to spherical micelles, elliptical micelles, rodlike micelles, vesicles, and a planar lipid bilayer. The geometrical parameters were optimized to fit the data. In each case, the model with the best chi-square value was chosen. The form factor of vesicles was derived from the form factor of a sphere

$$F_s(\rho, \rho_0, R, q) = \left( \frac{4}{3} \pi R^3 \right) (\rho - \rho_0) \times \left[ \frac{3 \sin(Rq) - Rq \cos(Rq)}{(Rq)^3} \right].$$

The form factor of a spherical micelle is then given by

$$F_v(\rho_H, \rho_0, R, D, q) = F_s(\rho_H, \rho_0, R, q) - F_s(\rho_H, \rho_c, R - D, q),$$

where  $\rho_H$ ,  $\rho_c$ , and  $\rho_0$  denote the electron densities of the lipid head group, the acyl chains, and the solvent, respectively, and where  $R$  is the radius of the micelles. The form factor of a vesicle consists of several shells,

$$F_v(\rho_s, \rho_0, R, D_1, D_2, q) = F_s(\rho_H, \rho_0, R, q) - F_s(\rho_H, \rho_c, R - D_1, q) + F_s(\rho_H, \rho_c, R + D_1 - D_2, q) - F_s(\rho_H, \rho_c, R + D - D_2, q).$$

$D_1$  and  $D_2$  are the thickness of the headgroup and the membrane thickness, respectively. Elliptical micelles were fitted with the anisotropic form factor

$$F^2(\rho_c, \rho_s, \rho_0, R_c, R_s) = \int_0^{\pi/2} [V_c(\rho_c - \rho_s)f(q, R_c) + (V_s + V_c) \times (\rho_s - \rho_0)f(q, R_s)]^2 \sin \Theta d\Theta,$$

$$f(q, R) = 3 \frac{\sin X - X \cos X}{X^3},$$

$$X = qR \sqrt{\sin^2 \Theta + \varepsilon^2 \cos^2 \Theta},$$

$$V = 4/3 \pi \varepsilon R^3,$$

where  $\varepsilon$  denotes the anisotropy factor.

## F. Continuous bleaching

The continuous bleaching method<sup>24,35,36</sup> was used to characterize the dynamics of the generated membranes. According to the theory of the method, the spatial intensity of a fluorescently labeled membrane is described by simultaneous photobleaching and replenishment of fluorescent molecules as they diffuse in two dimensions. Briefly, we continuously illuminate a defined region [with an approximate diameter of  $180\text{ }\mu\text{m}$ , viewed at  $63\times$  magnification, Figs. 7(a) and 7(e)] of the SLB, resulting in the bleaching of the Oregon green dye, whose fluorescence intensity at the center of the illuminated area,  $I_{d/2}$ , decays exponentially as a function of exposure time according to the equation

$$I_{d/2}(t) = I_0 e^{-B_0 t} = I_0 e^{-\tau},$$

where  $I_0$  is the initial fluorescence intensity at the center of the illuminated area,  $B_0$  is the bleaching rate, and  $\tau = B_0 t$  is the dimensionless time. Values for  $I_{d/2}$  and  $B_0$  were obtained as previously described.<sup>31</sup> When the fluorescence intensity at the center of the illuminated region approaches background fluorescence, the fluorescence intensity line profile [Figs. 7(c) and 7(g)] is fitted as a function of distance according to the equation

$$I_x = I_{x_0} \cosh \left[ \sqrt{\frac{B_0}{D}} \left( x - \frac{d}{2} \right) \right] + A'_2,$$

where  $I_{x_0}$  is the fluorescence intensity at the edge of the rim,  $D$  is the diffusion constant, and  $A'_2$  is a constant used in fitting. Diffusion constants are estimated from these fits [Figs. 7(c) and 7(f)]. SLBs formed by controlled deposition were compared to SLBs formed by vesicle fusion, as described by Zantl.<sup>37</sup>

## III. RESULTS AND DISCUSSION

### A. Determination of volume phases by SAXS

We used SAXS in order to determine the lyotropic mesophases of lipids through the whole spectra of water:isopropanol concentration ratios, as well as through different temperatures. Figure 2 shows the scattering intensities of the lipid, DMTAP, dissolved in water:isopropanol mixtures through increasing water concentration. In pure isopropanol, we find a sharp Bragg reflection at  $1.42\text{ nm}^{-1}$ , corresponding to a structural periodicity of 4.4 nm. The periodicity increases to 4.7 nm when the water content is increased to 15%. These values lie within the bounds of typical values for a lamellar phase. However, the detailed structure of the

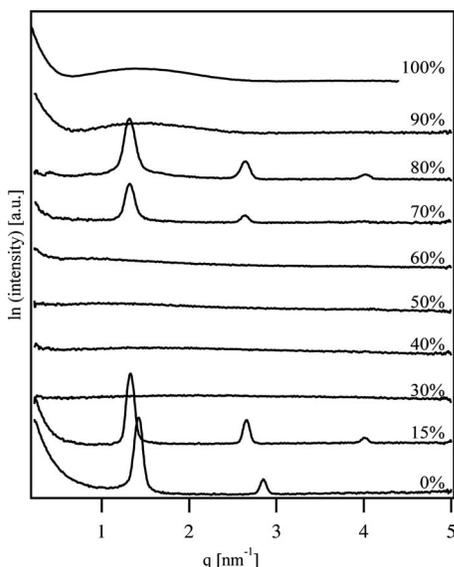


Fig. 2. DMTP scattering data for isopropanol-water solutions at 25 °C containing different amounts of water. The system goes from a lamellar phase in pure isopropanol and moves toward monomers and micelles, and eventually to a liposomal phase as the water content of the solvent increases. Data are plotted on a logarithmic scale.

lamellar phase in isopropanol might be distinct from the well-known lamellar phase of lipids dispersed in water.<sup>38</sup>

For 30% and 40% water contents (Fig. 2), no notable signals can be observed. These suggest a homogeneous scattering medium. A light bending that occurs maximally at  $1.8 \text{ nm}^{-1}$  corresponds to a distance of 3.5 nm, which is the characteristic correlation length for molecules at a concentration of 25 mg/ml. These features are both indicative of the

existence of lipid monomers. At 50% water content, a slightly increased intensity can be observed at  $1 \text{ nm}^{-1}$ , which may be indicative of a small fraction of micelles in solution. At 60%, the scattering profile exhibits a form factor which can be attributed to spherical micelles. The data were fitted to the form factor of a micelle in the range between 0.6 and  $5 \text{ nm}^{-1}$ , as described in the experimental section. Micelle-like structures are known to occur in binary isopropanol-water mixtures, without lipid, at 60% water content.<sup>38,39</sup> It is likely that in the ternary system, phospholipids will incorporate into these spherical aggregates. Our data are consistent with a polydisperse, Gaussian distributed system of spherical micelles with a radius of  $2.5 \text{ nm} \pm 1 \text{ nm}$  and a headgroup size of 0.2 nm. These data were derived from SAXS data assuming a core-shell structure as described in the experimental section.

At 70% and 80% water contents, Bragg reflections associated with a multilamellar phase were detected again, with a periodicity of 4.8 nm. At 90%, we find a form factor corresponding to dispersed unilamellar vesicles [Figs. 2 and 3(b), and as described in the experimental section]. In this case, the fitting analysis yields a range of possible fit parameters dependent on the selected size distributions, as well as on the electron density. However, values for the thickness of the lipid bilayer varied around a value of 4.5 nm. The data are consistent with an average vesicle radius ranging from 20 to 100 nm. The inaccuracy of the fit with respect to the vesicle radius is possibly indicative of high polydispersity.

The abundance of DMTP-isopropanol-water phases becomes even greater as the temperature is varied. As shown in Fig. 4(a), a region of inverted spherical micelles appears when the temperature is changed. Here, the radius of a mi-

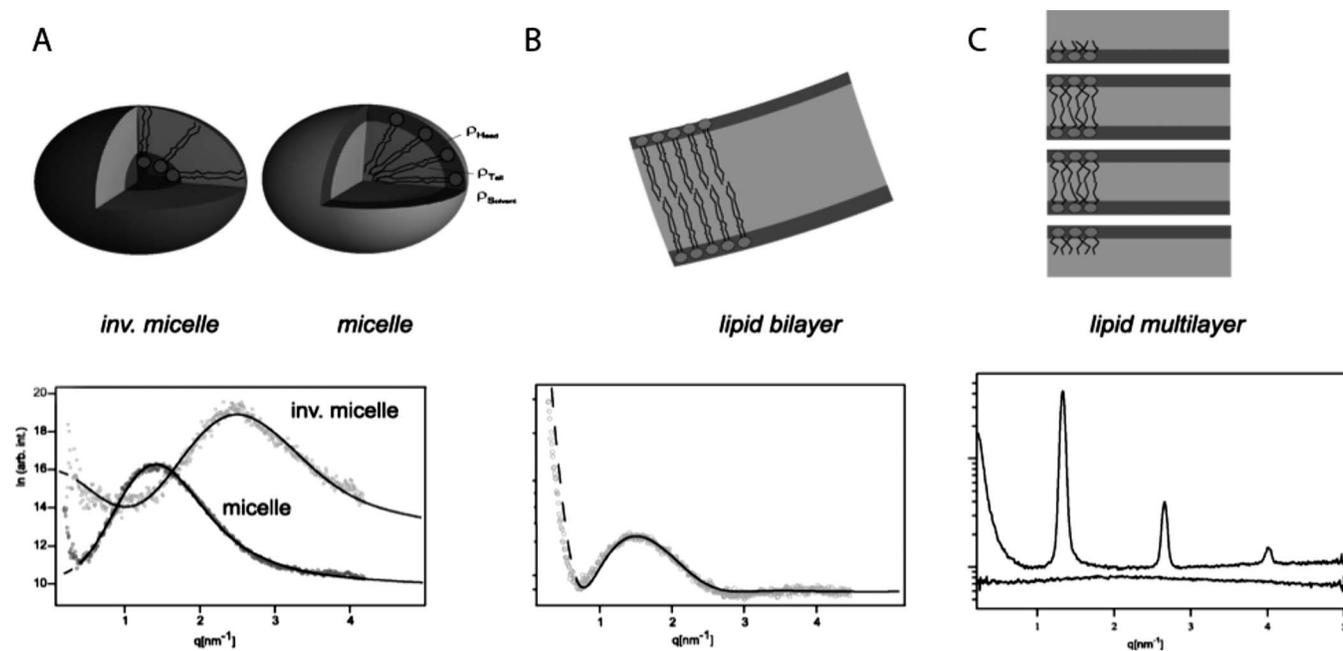


Fig. 3. Scattering data and corresponding form factor fits for various lipid structures in a DMTP solution, such as micelles and inverse micelles at 60% water content (a), lipid bilayers at 90% water content (b), and multilamellar phases at 15% water content (c). The monomeric phase, which has essentially no signal, is contrasted with the multilamellar phase (c).

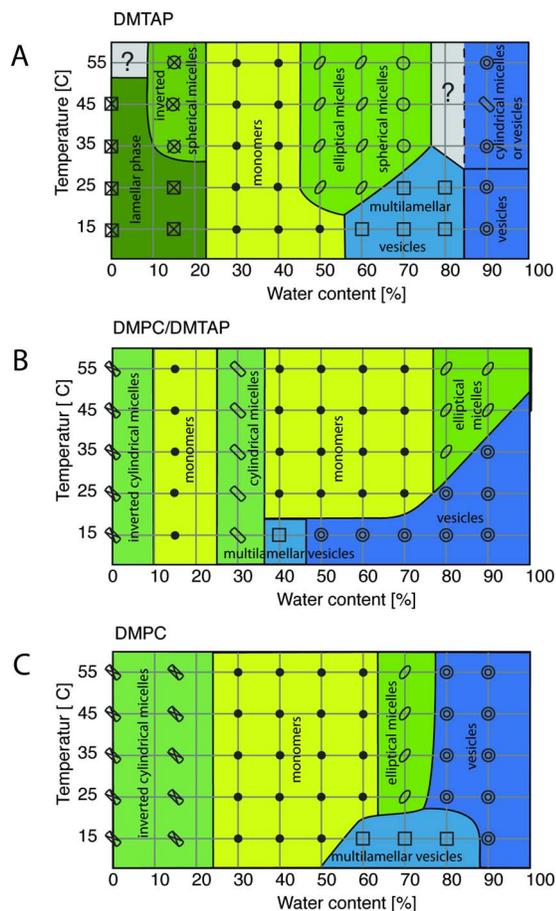


FIG. 4. (Color online) Phase diagrams for DMTAP (a), DMPC/DMTAP (b), and DMPC (c) with reference to varying isopropanol-water content and temperature. Lipids used in these measurements had a constant concentration of 25 mg/ml.

celle is around 2 nm, with a size distribution width of 0.1 nm. In Secs. III B–III D, we discuss phase behavior as a function of water content, temperature and lipid composition.

## B. Bulk phase diagrams

Complete phase diagrams of DMTAP, DMTAP/DMPC, and DMPC as a function of water-isopropanol content and temperature are shown in Fig. 4. All phase divisions have degrees of uncertainty of 5 °C and 5% water content since phase transitions were not measured along the full division; measurements were instead made in representative points in the phase diagram. Areas with a question mark could not be described using commonly used form factors. The phase behavior reveals a general phase order: With increasing water content, inverse micelles, monomer, micelles, and finally, vesicles can be found. This behavior is clearly represented by the phase diagram of DMPC. In all three phase diagrams, regions where the lamellar phase occurs can be seen at lower temperatures.

Upon the addition of DMTAP to DMPC, the phase behavior becomes more complex. The phase diagram for a 1:1 DMPC-DMTAP mixture [Fig. 4(b)] differs from the DMTAP diagram in terms of the absence of the lamellar phase at low

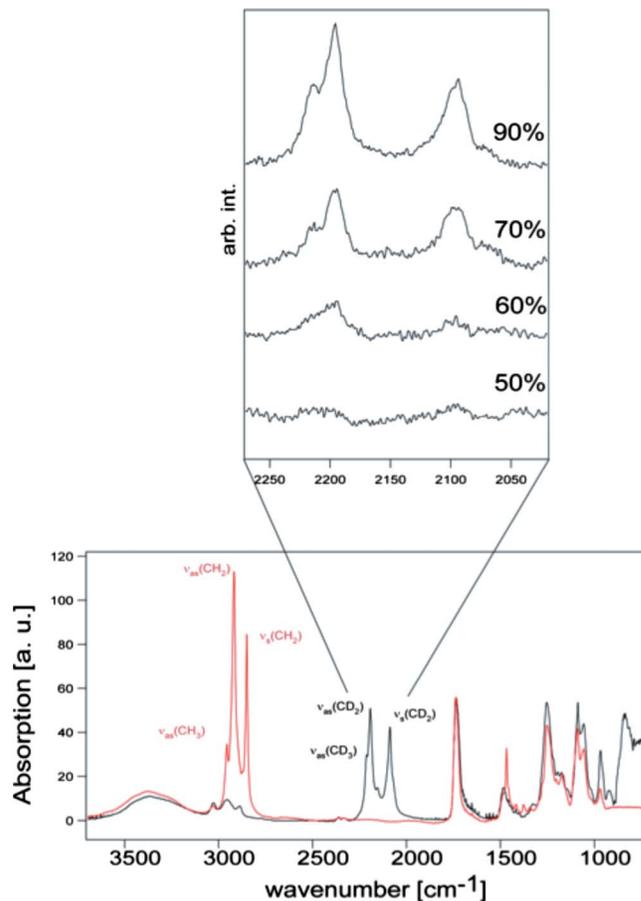


FIG. 5. (Color online) Infrared spectra showing the C–H and C–D stretching vibrations of DMPC (red line) and d54-DMTAP (black line, inset). The C–D stretching vibrations in the ATR mode are shown in the expansion as a function of increasing water content.

water content. Furthermore, higher water content is required for the micelle phase to occur. In pure isopropanol, the lipid mixture exists as inverted cylindrical micelles. On reaching a water content of 15%, a transition to monomer phase is achieved, which could in turn undergo a transition to cylindrical micelles once the water content is increased to 30%. It is remarkable that vesicles at high water content but elevated temperatures are no longer stable. In the same regime, DMTAP exhibits phases that were not unambiguously determinable. Since we are interested how one can enter vesicle phases, it is remarkable that DMTAP/DMPC exhibits a temperature-induced transition from monomer to vesicles at water content of  $\geq 50\%$ .

## C. Lipid surface coverage measured by ATR-FTIR

In this section, we discuss the adsorption of lipids from the bulk phase described above to silicon oxide as a function of water content. To distinguish the lipid from isopropanol, we used phospholipids with deuterated alacyl chains, which allow the determination of the degree of adsorption through specific infrared adsorption bands from ATR-FTIR. Figure 5 shows a bulk infrared spectrum of DMPC and deuterated d54-DMPC indicating the position of the C–H and C–D

stretching bands. The position of the C–H bands overlaps with the adsorption bands of isopropanol and cannot be used. The C–D stretching bands of the lipid, however, can be detected in the ATR mode, despite a broad background signal arising from water and isopropanol in bulk solution. Consequently, reference measurements were made for each sample, and a corrected spectrum is obtained, as described in the experimental section. The expansion in Fig. 5 shows typical results from a series of ATR-FTIR measurements.

Surface occupancy is proportional to the intensity of the alkyl bands. To determine this intensity, the spectra were matched against a sum of six Gaussian curves, which match the shape of the adsorption band fingerprint. In the fit, the amplitude of this set of Gaussians was taken as a measure of the degree of lipid coverage. The position and width of the set of Gaussians were independently determined from reference measurements, where higher lipid concentrations were used. Figure 6 shows the isothermal adsorption of DMTAP. Results are normalized to 1.0 at 100% water content, corresponding to the surface coverage obtained from vesicle fusion, i.e., a single lipid bilayer. An arc-tan fit, which serves as a visual guide, is overlaid across the adsorption isotherm as a function of water content. For DMTAP, the point at which half-surface coverage can be observed appears to be at 67% water content.

In the case of DMPC, a significantly shallower isotherm is observed, which does not have a defined transition point. Half-maximum coverage is found at approximately 40% water content; finite adsorption is observable even in pure isopropanol (0% water). In the case of the equimolar lipid mixtures of DMPC-DMTAP, the transition occurs between 63% water content. In comparison with pure DMTAP and pure DMPC, this approximates a behavior that is between the two. Figure 6(c) also indicates the percentages of DMPC and DMTAP separately. These data can be obtained if only one of the lipid components is deuterated. More interestingly, there is always a larger amount of DMPC than DMTAP adsorbed to the oxidized silicon surface. Infrared spectroscopy results indicate that the addition of water causes lipids dissolved in isopropanol to adhere to the hydrophilic silicon oxide surface. These ATR-FTIR results, however, do not reveal details of the quality of the lipid coating, specifically its homogeneity and its lateral diffusion.

#### D. Adsorption behavior of a thermotropic system

SAXS results imply that lipid adsorption can be induced not only through changes in solvent composition but through temperature changes as well. This strong temperature-dependent behavior is manifested through an increased degree of coverage for lower temperatures. If DMTAP in a 90% water:isopropanol solution is heated to 45 °C, all the lipids desorb. Upon cooling, a lipid bilayer is re-adsorbed, with increased coverage compared to before heating. In order to confirm if this effect on coverage is characteristic of the system or if it is an artifact of the prolonged equilibration period, DMTAP samples at different water contents are warmed from 10 to 50 °C and then cooled to 10 °C again.

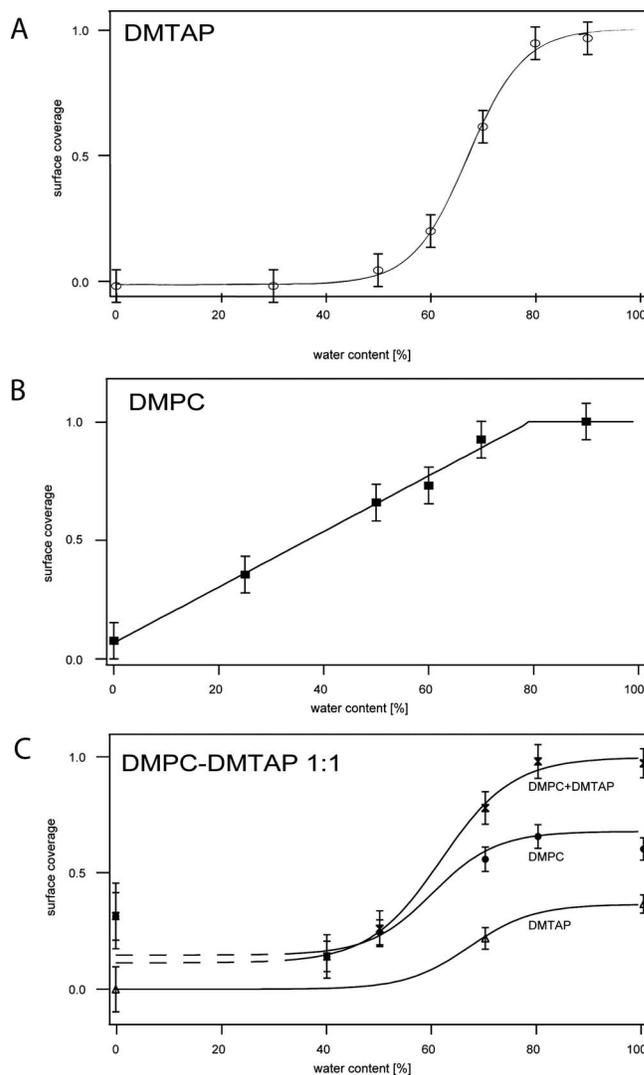


Fig. 6. Isothermal adsorption curves for d54-DMTAP (a), d54-DMPC (b), and d54-DMPC/d54-DMTAP (c) at 28 °C under varying water content. C–D stretching vibration intensities were normalized, and correspond to the lipid concentration on the silicon surface. The error bars are taken from the fits of the spectra; for multiple measurements, errors were calculated according to the rules of error propagation.

Surface coverage was checked at intervals of 10 °C during the process of heating, as well as after cooling. Results confirmed the initial observation on the temperature-dependence of coverage (data not shown). In terms of the phase changes, our hypothesis is that increasing the temperature at 90% water content results in transitions from the vesicle/bilayer phase to a mixed micelle and vesicle phase [Fig. 4(a)], and it is possible that the micelles generated impinge on the surface on cooling, and is responsible for the increased coverage observed (Fig. 1). These observations are reasonably consistent with previous reports correlating surface coverage and temperature, where an annealing and cooling cycle results in a saturated lipid film.<sup>40</sup>

#### E. Diffusion measurements

In order to evaluate the mobility and homogeneity of the membranes on a microscopic-length scale, continuous

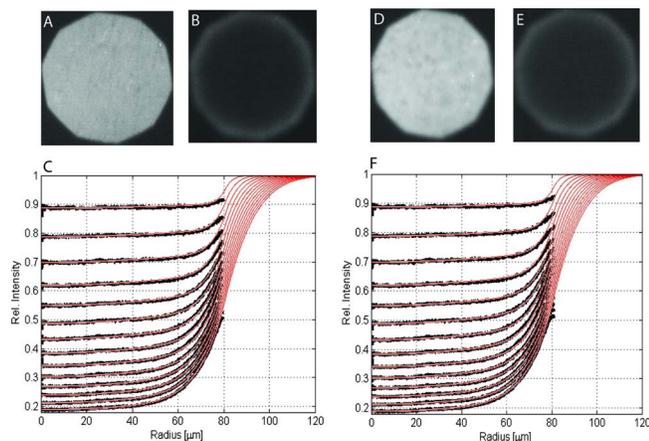


Fig. 7. (Color online) Continuous bleaching showing lipid diffusion in Oregon green-labeled DMPC SLBs prepared using the solvent-exchange method [(a) and (b)] and vesicle fusion [(d) and (e)]. Average background-corrected intensity profiles from the edge to the center of a circular selection made from the decagonal sample were used in obtaining the bleaching rate, which was subsequently used in calculating the diffusion constant based on fitting these profiles by the equations specified in the methods section [(c) and (f)]. The diffusion constant obtained for both SLBs was  $2.6 \pm 0.3 \mu\text{m}^2/\text{s}$ . Micrographs (a) and (c) were taken at  $t=0$ , and micrographs (b) and (e) were taken at  $t=1$  min. For the experiments, illuminated regions with an approximate diameter of  $180 \mu\text{m}$  were used. Continuous bleaching experiments were also performed for longer-chain SOPC membranes, and were found to have a diffusion constant of  $2.3 \pm 0.4 \mu\text{m}^2/\text{s}$  (data not shown).

bleaching experiments were used. We directly compared supported lipid bilayers prepared using the solvent exchange method to those prepared by vesicle fusion. Lipid bilayer deposition onto a glass surface by solvent exchange was achieved using a flow-through sample chamber filled with lipid in an 8:2 water:isopropanol mixture. The water content was subsequently increased by the slow titration of the lipid solution. These were compared to membranes generated through small unilamellar vesicle fusion according to standard protocols.<sup>12,37,41,42</sup> A visual evaluation of the membranes produced by solvent exchange and vesicle fusion shows that both are homogeneous [Figs. 7(a) and 7(d)]. The samples were subsequently bleached continuously to evaluate its fluidity. In continuous bleaching, illumination is concentrated on a circular spot of the membrane using a field aperture. If a membrane is fluid, the fluorescence signal at the center of the illuminated area decays exponentially over time, while a fluorescence rim forms at its periphery. The latter results from the continuous diffusion of unbleached, fluorescently labeled lipids from the surrounding nonilluminated area. Both samples exhibit these characteristics of fluid membranes [Figs. 7(b) and 7(e)]. Evaluation of the diffusion data of the monolayer in contact with the surface indicate that membranes formed through solvent exchange have diffusion constants of  $2.6 \pm 0.3 \mu\text{m}^2 \text{s}^{-1}$  comparable with those obtained by vesicle fusion protocol, which is found to be also  $2.6 \mu\text{m}^2 \text{s}^{-1}$  within experimental accuracy (Fig. 7) (for details, please refer to the experimental section). The measured diffusion constant consistent with the value  $2.0 \pm 0.4 \mu\text{m}^2 \text{s}^{-1}$  obtained for a DMPC bilayer prepared by

the vesicle fusion technique.<sup>41,42</sup> For comparison, the reader is reminded that in a free lipid bilayer, the diffusion constant of DMPC lies at about  $0.5\text{--}5 \mu\text{m}^2 \text{s}^{-1}$  at room temperature.<sup>41</sup> For this protocol, the washing steps to remove excess alcohol are important; otherwise, the presence of a significant amount of residual alcohol would tend to increase membrane fluidity as a result of its H-bonding with the hydrocarbon chains, which then leads to a decrease its conformation order.<sup>43</sup>

## F. Conclusion

We have investigated the phase behavior of the ternary system comprised of lipid, isopropanol, and water as a function of water content, temperature, and lipid head group. From SAXS and ATR-FTIR measurements, the bulk phase behavior was found to correlate with the amount of surface coverage. With increasing water content, the system undergoes phase transitions through the bulk phases: inverted micellar-monomeric-micellar-vesicles. At the same time, the coverage of lipid on the silicon oxide surface of the ATR crystal increases monotonically, reaching the saturating level of a single lipid bilayer coating. There are remarkable differences between the behavior of DMTAP and DMPC, with DMTAP exhibiting a steep increase in coverage as opposed to DMPC, which shows gradual deposition with increasing water content. This indicates that DMPC has a higher affinity for the solid than the cationic DMTAP, despite the fact that the solid is negatively charged. We also argue that within the transition regime from micelles to vesicles, the formation of solid supported bilayer is energetically favored over vesicle formation in the bulk phase due to the fact that surface interactions enhance the relative concentration of lipids at the surface. From this, we can define a protocol for depositing a lipid bilayer, which involves the incubation of a surface with a ternary mixture below this critical micelle-to-vesicle transition and gradually increasing the water content. A theoretical basis for this can be taken from the thermodynamics of alcohols interacting with lipid membranes.<sup>44,45</sup> We assume that the lipid aggregates at the solid surface, either monomers or micelles, are in thermodynamic equilibrium with the bulk phase. As the solvent is varied toward a pure water phase, the critical micelle concentration for the lipid is lowered and the lipids bound on the surface close to form a bilayer. This is in contrast to the vesicle fusion method, where lipid deposition takes place via vesicle rupture, an event that is associated with a high energy barrier. Additionally, lipids are required to spread and anneal at the surfaces after vesicle fusion. Hence, in special cases, where vesicle fusion or spreading is limited, as the case is for hydrophobic or nanostructured surfaces, the solvent exchange method might be advantageous for the formation of a supported bilayer.<sup>24</sup> Compared to vesicle fusion, this method has a clear shortcoming in that incorporation of proteins and precise control over composition are not guaranteed, and would require further investigation. For cases where a one-component lipid coating is required, however, the solvent-exchange protocol has advantages in terms of its applicability to complex structures, such as nanocontainers,

microfluidic devices with extended networks, or solid particles. In particular, the surface functionalization of microfluidic channels with solid supported bilayers, which aims to reduce protein-wall interactions, is an example where the solvent exchange technique can be applied.<sup>46</sup> For this kind of technology-driven applications, the ease of preparation is of primary importance, and the method presented here might have most impact.

## ACKNOWLEDGMENTS

This work was carried out in the former biophysics laboratory of Erich Sackmann. J.R. and A.H. are grateful for the candid mentoring and stimulating discussions with him. M.P.D. acknowledges the Deutscher Akademischer Austausch Dienst (DAAD) for her Ph.D. scholarship.

- <sup>1</sup>E. Sackmann, *Science* **271**, 43 (1996).
- <sup>2</sup>M. Tanaka and E. Sackmann, *Nature (London)* **437**, 656 (2005).
- <sup>3</sup>J. Spinke, J. Yang, H. Wolf, H. Ringsdorf, and W. Knoll, *Biophys. J.* **63**, 1667 (1992).
- <sup>4</sup>Y. Chan and S. Boxer, *Curr. Opin. Chem. Biol.* **11**, 581 (2007).
- <sup>5</sup>B. Sani and A. Parikh, *Annu. Rev. Phys. Chem.* **59**, 411 (2008).
- <sup>6</sup>J. Groves, N. Ulman, and S. Boxer, *Science* **275**, 651 (1997).
- <sup>7</sup>C. Steinem, A. Janshoff, W. P. Ulrich, M. Sieber, and H.-J. Galla, *Biochim. Biophys. Acta* **1279**, 169 (1996).
- <sup>8</sup>S. Lenhart, P. Sun, Y. Wang, and H. Fuchs, *Small* **3**, 71 (2007).
- <sup>9</sup>S. Gritsch, P. Nollert, F. Jähnig, and E. Sackmann, *Langmuir* **14**, 3118 (1998).
- <sup>10</sup>D. Olson, J. Johnson, P. Patel, E. S. G. Shaqfeh, S. G. Boxer, and G. G. Fuller, *Langmuir* **17**, 7396 (2001).
- <sup>11</sup>V. Kahl, M. Hennig, B. Maier, and J. Rädler, *Electrophoresis* **30**, 1276 (2009).
- <sup>12</sup>M. Hochrein, J. Leierseder, L. Golubovic, and J. Rädler, *Phys. Rev. Lett.* **96**, 038103 (2006).
- <sup>13</sup>J. Groves, *Curr. Opin. Chem. Biol.* **10**, 544 (2006).
- <sup>14</sup>I. Czolkos, Y. Erkan, P. Dommersnes, A. Jesorka, and O. Orwar, *Nano Lett.* **7**, 1980 (2007).
- <sup>15</sup>L. Tamm and H. McConnell, *Biophys. J.* **47**, 105 (1985).
- <sup>16</sup>H. McConnell, T. Watts, R. Weis, and A. Brian, *Biochim. Biophys. Acta* **864**, 95 (1986).
- <sup>17</sup>E. Kalb, S. Frey, and L. Tamm, *Biochim. Biophys. Acta* **1103**, 307 (1992).
- <sup>18</sup>J. Nissen, S. Gritsch, G. Wiegand, and J. Rädler, *Eur. Phys. J. B* **10**, 335 (1999).
- <sup>19</sup>J. Raedler, H. Strey, and E. Sackmann, *Langmuir* **11**, 4539 (1995).
- <sup>20</sup>U. Mennicke and T. Salditt, *Langmuir* **18**, 8172 (2002).
- <sup>21</sup>C. J. Brinker, Y. Lu, A. Sellinger, and H. Fan, *Adv. Mater. (Weinheim, Ger.)* **11**, 579 (1999).
- <sup>22</sup>E. Sackmann and M. Tanaka, *Trends Biotechnol.* **18**, 58 (2000).
- <sup>23</sup>F. Giess, M. G. Friedrich, J. Heberle, R. L. Naumann, and W. Knoll, *Biophys. J.* **87**, 3213 (2004).
- <sup>24</sup>M. Hochrein, C. Reich, B. Krause, J. Radler, and B. Nickel, *Langmuir* **22**, 538 (2006).
- <sup>25</sup>T. Bayerl and M. Bloom, *Biophys. J.* **58**, 357 (1990).
- <sup>26</sup>F. Tiberg, I. Harwigsson, and M. Malmsten, *Eur. Biophys. J.* **29**, 196 (2000).
- <sup>27</sup>F. Szoka and D. Papahadjopoulos, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 4194 (1978).
- <sup>28</sup>J. Strassburger and I. Smith, *Appl. Spectrosc.* **33**, 283 (1979).
- <sup>29</sup>E. Goormaghtigh and J. Ruyschaert, *Spectrochim. Acta, Part A* **50**, 2137 (1994).
- <sup>30</sup>J. Kalus and M. C. Gerstenberg, *Phys. Scr.* **T49B**, 629 (1993).
- <sup>31</sup>N. P. Franks and Y. K. Levine, *Mol. Biol. Biochem. Biophys.* **31**, 437 (1981).
- <sup>32</sup>G. Pabst, M. Rappolt, H. Amenitsch, and P. Laggner, *Phys. Rev. E* **62**, 4000 (2000).
- <sup>33</sup>O. Glatter and O. Kratky, *Small Angle X-Ray Scattering* (Academic, London, 1982).
- <sup>34</sup>J. S. Higgins and H. C. Benoît, *Polymers and Neutron Scattering* (Oxford Science, New York, 1994).
- <sup>35</sup>C. Dietrich, R. Merkel, and R. Tampé, *Biophys. J.* **72**, 1701 (1997).
- <sup>36</sup>M. Horton, C. Reich, A. Gast, J. Rädler, and B. Nickel, *Langmuir* **23**, 6263 (2007).
- <sup>37</sup>R. Zantl, *Diplomarbeit* (Technische Universität, München, 1997).
- <sup>38</sup>F. Ricoul, M. Dubois, T. Zemb, and D. Plusquellec, *Eur. Phys. J. B* **4**, 333 (1998).
- <sup>39</sup>M. D'Angelo, G. Onori, and A. Santucci, *J. Chem. Phys.* **100**, 3107 (1994).
- <sup>40</sup>E. Reimhult, M. Zeh, F. Hk, and B. Kasemo, *Langmuir* **22**, 3313 (2006).
- <sup>41</sup>C. K. Yee, M. L. Amweg, and A. N. Parikh, *J. Am. Chem. Soc.* **126**, 13962 (2004).
- <sup>42</sup>M. P. Goertz, B. L. Stottrup, J. E. Houston, and X. Y. Zhu, *J. Phys. Chem. B* **113**, 9335 (2009).
- <sup>43</sup>M. Patra, E. Salonen, E. Terama, I. Vattulainen, R. Faller, B. W. Lee, J. Holopainen, and M. Karttunen, *Biophys. J.* **90**, 1121 (2006).
- <sup>44</sup>P. Westh and C. Trandum, *Biochim. Biophys. Acta* **1421**, 261 (1999).
- <sup>45</sup>J. Schellman, *Biophys. Chem.* **37**, 121 (1990).
- <sup>46</sup>R. Galneder, V. Kahl, A. Arbuzova, M. Rebecchi, J. O. Rädler, and S. McLaughlin, *Biophys. J.* **80**, 2298 (2001).

# Appendix E

## Original publication P5

Maria Pamela Dobay, Akos Dobay, Johnrob Bantang, Eduardo Mendoza

**How many trimers? Modeling influenza virus fusion yields a minimum aggregate size of six trimers, three of which are fusogenic**

*Mol. BioSyst.* **7**: 2741 - 2749, 2011.

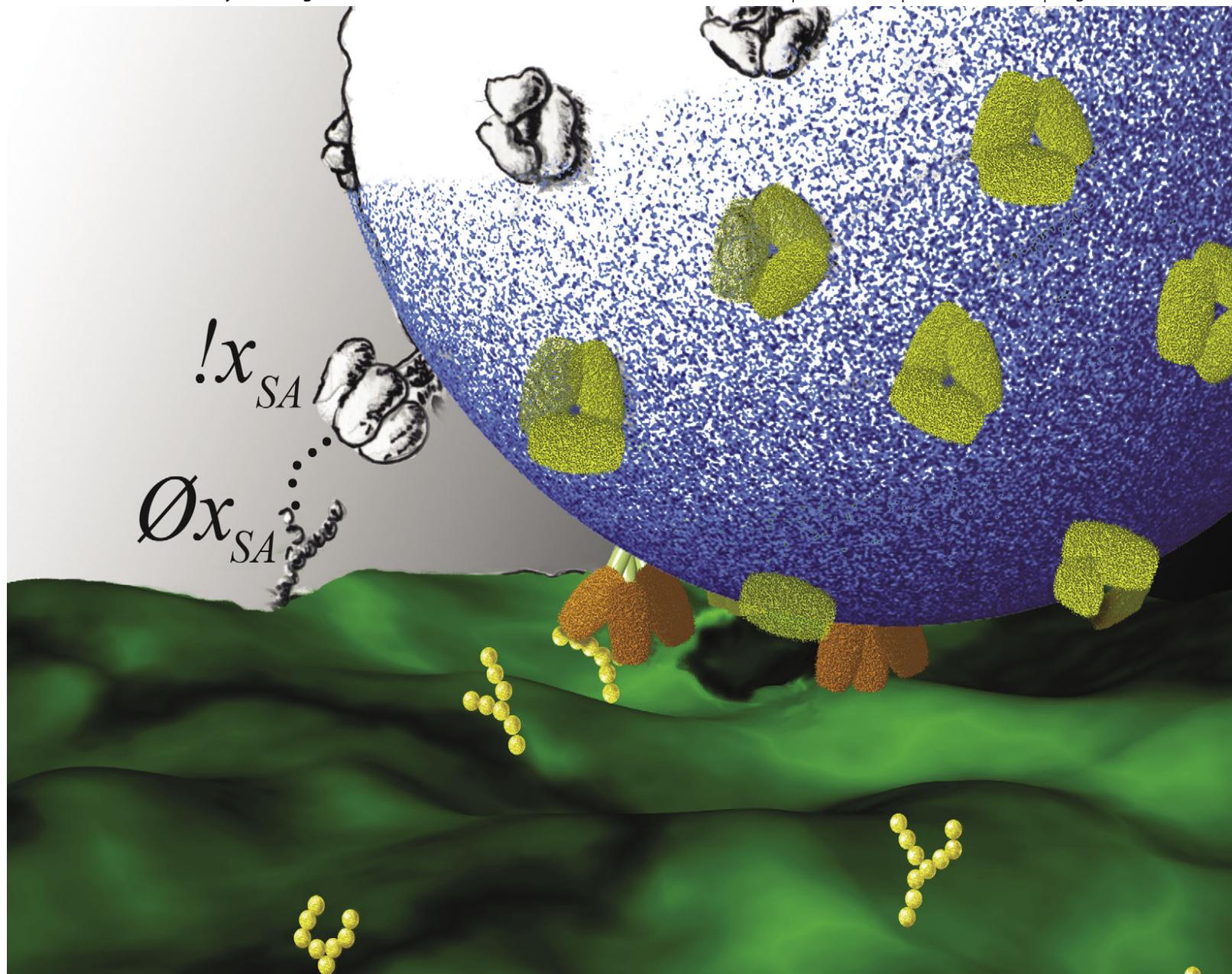


# Molecular BioSystems

Indexed in  
**MEDLINE!**

www.molecularbiosystems.org

Volume 7 | Number 10 | 1 October 2011 | Pages 2723–2910



Themed issue: Computational Biology

ISSN 1742-206X

RSC Publishing

**PAPER**

Maria Pamela Dobay *et al.*

How many trimers? Modeling influenza virus fusion yields a minimum aggregate size of six trimers, three of which are fusogenic



1742-206X(2011)7:10;1-3

## How many trimers? Modeling influenza virus fusion yields a minimum aggregate size of six trimers, three of which are fusogenic†‡

Maria Pamela Dobay,<sup>§\*a</sup> Akos Dobay,<sup>bc</sup> Johnrob Bantang<sup>d</sup> and Eduardo Mendoza<sup>ae</sup>

Received 13th February 2011, Accepted 16th June 2011

DOI: 10.1039/c1mb05060e

Conflicting reports in leading journals have indicated the minimum number of influenza hemagglutinin (HA) trimers required for fusion to be between one and eight. Interestingly, the data in these reports are either almost identical, or can be transformed to be directly comparable. Different statistical or phenomenological models, however, were used to analyze these data, resulting in the varied interpretations. In an attempt to resolve this contradiction, we use PABM, a brane calculus we recently introduced, enabling an algorithmic systems biology approach that allows the problem to be modeled in a manner following a biological logic. Since a scalable PABM executor is still under development, we sufficiently simplified the fusion model and analyzed it using the model checker, PRISM. We validated the model against older HA-expressing cell-to-cell fusion data using the same parameters with the exception of three, namely HA and sialic acid (SA) surface densities and the aggregation rate, which were expected to be different as a result of the difference in the experimental setup. Results are consistent with the interpretation that a minimum aggregate size of six HA trimers, of which three undergo a conformational change to become fusogenic, is required for fusion. Of these three, two are free, while one is bound. Finally, we determined the effects of varying the SA surface density and showed that only a limited range of densities permit fusion. Our results demonstrate the potential of modeling in providing more precise interpretations of data.

### 1 Introduction

Membrane fusion is one of the most fundamental biological processes exhibiting mechanistic similarities across its different forms, from viral and bacterial entry to intracellular fusion.<sup>1,2</sup> Most of what is known regarding viral fusion are derived from structural and mechanistic studies of influenza virus fusion, as mediated by hemagglutinin (HA). Hemagglutinin is a trimeric protein anchored to the viral membrane *via* its C-terminal domain.<sup>3</sup> HA binds host cells through sialic acid (SA)-capped

proteins, which are particularly abundant in the respiratory tract, as well as in red blood cells.<sup>4</sup> Binding triggers internalization into an endosome, the acidification of which causes HA trimers to aggregate and to undergo a conformational change that extends its N-terminal fusion peptides and causes it to fuse with the endosome membrane, and release its contents into the host cell cytosol.<sup>3</sup> *In vitro* experiments have expanded the view of HA-mediated fusion through the definition of intermediates between the conformational change step and content mixing. These include the generation of the first fusion pore (FP), through which ions can pass between the virus and target membranes; the lipid channel (LC), which permits the lipids to mix between the two membranes; and the formation of the fusion site (FS), which allows content mixing<sup>5</sup> (Fig. 1).

In an attempt to characterize membrane fusion and its intermediates better, several groups have designed experiments to determine the minimum requirements for the formation of a fusion pore. Viruses or virus-like systems, which have been evolved to efficiently form such pores on the endosome membrane through proteins such as HA, are consequently ideal for such studies. Knowing these minimum fusion requirements is also of interest in artificial gene and drug therapy, where efficient endosomal escape remains one of the main problems. We are particularly interested in designing HA-decorated vectors for gene and drug delivery. These vectors,

<sup>a</sup> Faculty of Physics and Center for NanoScience, Ludwig-Maximilians-University, Geschwister-Scholl-Platz 1, Munich, Germany. E-mail: maria.pamela.david@physik.uni-muenchen.de; Fax: +49-(0)89-2180-3182; Tel: +49-(0)89-2180-1453

<sup>b</sup> Institute of Evolutionary Biology and Environmental Studies, University of Zurich, Winterthurststrasse 190, 8057 Zurich, Switzerland

<sup>c</sup> Swiss Institute of Bioinformatics, Quartier Sorge Bâtiment Génopode, CH-1015 Lausanne, Switzerland

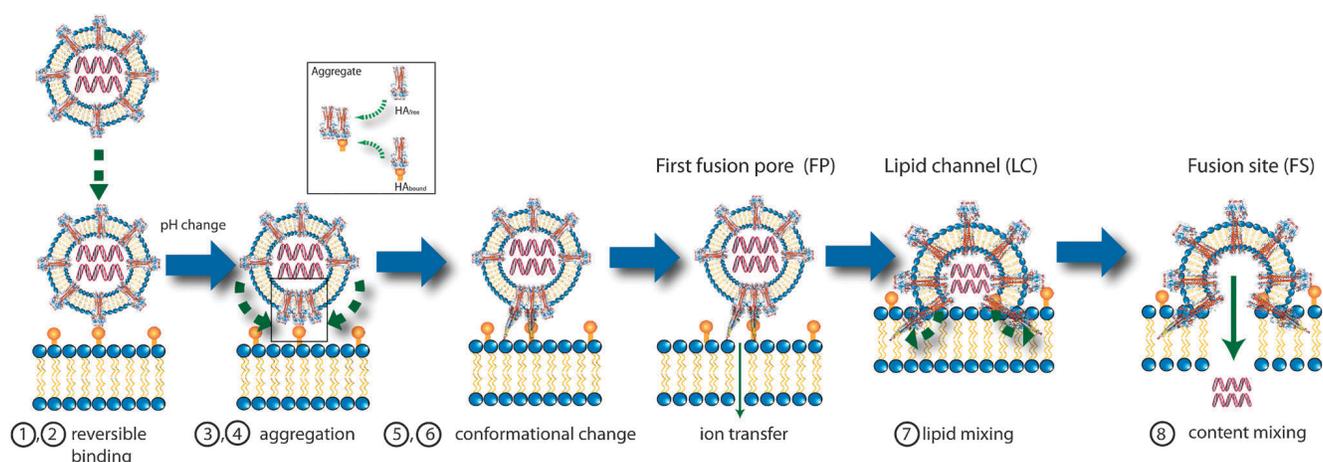
<sup>d</sup> National Institute of Physics, University of the Philippines, Diliman, Quezon City, Philippines

<sup>e</sup> Institute of Mathematics, University of the Philippines, Diliman, Quezon City, Philippines

† Published as part of a Molecular BioSystems themed issue on Computational Biology: Guest Editor Michael Blinov.

‡ Electronic supplementary information (ESI) available. See DOI: 10.1039/c1mb05060e

§ As obtained by Bentz.<sup>5</sup>



**Fig. 1** Steps in HA-mediated viral fusion *in vitro*. An influenza virus binds to sialic acid-capped receptors of the cell through the HA trimers at its surface. Changing the pH triggers the aggregation of other HA trimers at the contact site, as well as a conformational change in a subset of these trimers to form a fusion pore that allows ion exchange between the virus and target. Note that the aggregate can be comprised of both bound and unbound HA trimers (inset). Each of the steps are numbered to correspond to reactions in Section 2.2.

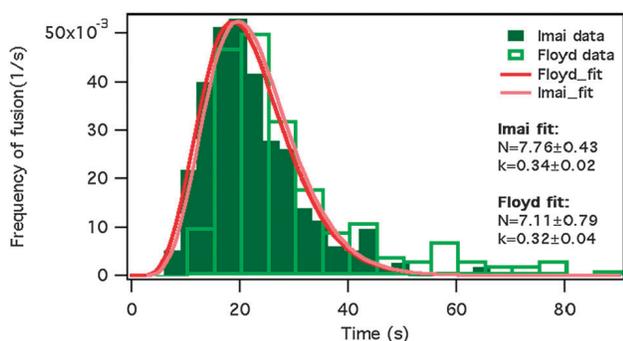
**Table 1** Minimum number of hemagglutinin trimers required for fusion as a function of experimental and statistical methods

| Paper  | $\omega$ | $q$ | Virus strain(s)/ cell line(s) | Fusion partner                | HA/contact area                      | SA/contact area                       | Detection method                 | Fitting and statistical methods                                 | Observed step |
|--|----------|-----|-------------------------------|-------------------------------|--------------------------------------|---------------------------------------|----------------------------------|---|---------------|
| Melikyan <i>et al.</i> <sup>12</sup>         | 8*       | n/a | HAB2, GP4f                    | Planar bilayer with fused RBC | $61 \times 10^3$ – $95 \times 10^3$  | $1.4 \times 10^6$ – $7.2 \times 10^6$ | Time-resolved admittance         | Exponential fit   | FP            |
| Blumenthal <i>et al.</i> <sup>13</sup>       | 6        | n/a | GP4f                          | RBC                           | $61 \times 10^3$                     | $1.4 \times 10^6$ – $7.2 \times 10^6$ | VFM                              | Empirical equation based on pore-opening kinetics               | LC, FS        |
| Danieli <i>et al.</i> <sup>8</sup>           | n/a      | 3   | HAB2, GP4f, gp4/6             | RBC                           | $37 \times 10^3$ – $479 \times 10^3$ | $1.4 \times 10^6$ – $7.2 \times 10^6$ | Spectrometry (bulk)              | Hill fit  | LC            |
| Guenther-Ausborn <i>et al.</i> <sup>11</sup> | n/a      | 1   | X-47, A Shangdong             | RBC                           | 20–30                                | 74–372                                | Resonance energy transfer (bulk) | Modified Hill fit   | LC, FS        |
| Imai <i>et al.</i> <sup>10</sup>             | n/a      | 1   | PR/8/1924                     | RBC                           | 20–30                                | 74–372                                | VFM                              | Log–log plots based on HA surface density and fusion rates      | LC            |
| Floyd <i>et al.</i> <sup>9</sup>             | n/a      | 3   | X-31                          | Planar bilayer with pure GD1a | 20–30                                | 74–372                                | VFM                              | $\Gamma$ -fit of frequency vs. time distribution of fusion data | FP/LC, FS     |

which could be even smaller than the viruses itself, should contain at least the minimum number of fusogenic units, but the least number of HAs that would still permit fusion, to reduce their potential immunogenicity.<sup>6</sup> In the case of HA, these minimum requirements refer to the number of aggregated trimers that comprise the fusion pore,  $\omega$ , which may be comprised of both HA bound to SA ( $HA_{\text{bound}}$ ) and free HA ( $HA_{\text{free}}$ ). A subset of  $\omega$ ,  $q$ , undergoes an acidification-mediated conformational change to a final fusogenic form.<sup>¶</sup> However, despite the fact that the experiments were directed towards the description of a single phenomenon, the results and interpretations derived from these vary. Table 1 summarizes results from different experimental groups, together with information on the experimental setup used, as well as the statistical

methods or phenomenological models used in data analysis, when applicable. Fusion intermediates that are possible to observe with each setup are also indicated. Typically, FP can be observed through conductivity measurements, while LC and FS are typically observed using video fluorescence microscopy (VFM).<sup>7</sup> The results describe the process as either being a cooperative<sup>8,9</sup> or a non-cooperative process.<sup>10,11</sup> Interestingly, a closer analysis of some of the data sets reveals that the results themselves are not so much varied as the analysis. For instance, a superimposition of fusion data from Imai and Floyd, which use comparable experimental setups, reveals that the experimental results are almost identical (Fig. 2). However, the conclusions of the studies are different, with Floyd and his co-authors supporting the idea of  $q = 3$  on the basis of a  $\Gamma$ -fit of their data, whereas Imai used an additional set of experiments involving variable fusogenic HA surface densities as the basis for a conclusion of  $q = 1$ .<sup>9,10</sup>

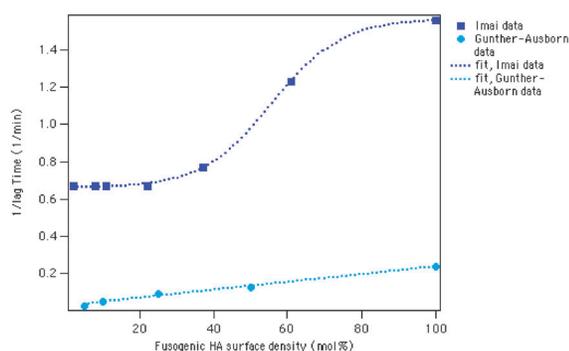
¶ The convention of using the notations  $\omega$  and  $q$  was taken from Bentz.<sup>5</sup>



**Fig. 2** Superposition of the data reported by Imai *et al.*<sup>10</sup> and Floyd *et al.*<sup>9</sup> indicates that the experimental results are almost identical. Nonetheless, the groups used different statistical methods for analysis. Fitting the Imai data set with an approximation of a  $\Gamma$  function yields a result of the same magnitude as in Floyd *et al.*; if this function is used as a reference for a conclusion, the data of Imai can be interpreted to support a conclusion of  $q = 3$ .

Consequently, if the Imai data had been interpreted on the basis of a  $\Gamma$ -fit, as the case was in Floyd, then it would have been in support of  $q = 3$ .

In contrast, there are also some studies where a similar experimental design was used, and that have arrived at the same conclusion that  $q = 1$ . Nonetheless, the experimental data obtained from these studies were very different, and were likewise analyzed using different techniques (Fig. 3).<sup>10,11</sup> In Fig. 3, the lag time is defined as the time interval between the exposure to low pH and the onset of fusion, and Gunther-Ausborn *et al.* postulated that the reciprocal of the lag time, as well as the initial rate of fusion are directly proportional to the surface density of fusion-competent HA trimers.<sup>11</sup> Obtaining a linear relationship consequently implies that the reaction is first order (*viz.*  $q = 1$ ) with respect to HA, whereas a non-linear relationship would imply  $q > 1$ . If the Imai data had been interpreted based on this method as shown in Fig. 3, then



**Fig. 3** Superposition of the data reported by Imai *et al.*<sup>10</sup> and Gunther-Ausborn *et al.*<sup>11</sup> using a  $1/\text{lag time vs. fusogenic HA surface density}$  plot. Gunther-Ausborn *et al.* define the lag time as the interval between sample exposure to low pH and the onset of fusion; they postulated that the relationship between the reciprocal lag time with the fusogenic HA surface density gives the order of the reaction with respect to HA. Their results yield a linear relationship, supporting  $q = 1$ . If the same analysis had been used on the results of Imai *et al.*, they would have obtained a nonlinear curve that supports  $q \neq 1$ , instead of  $q = 1$ .

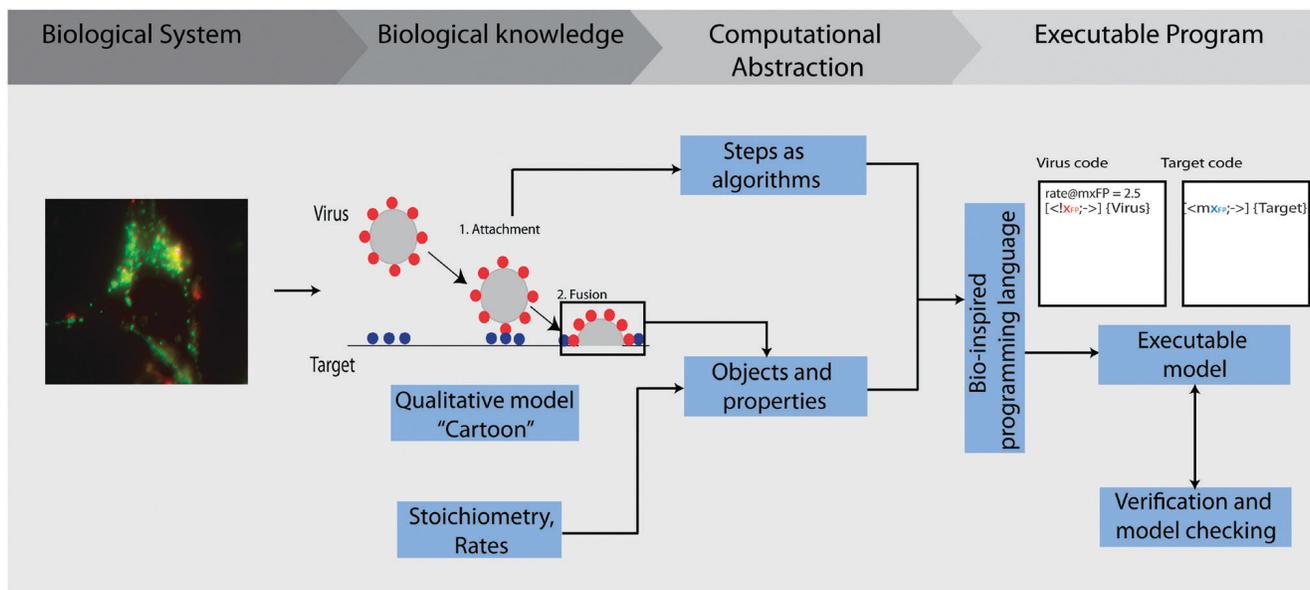
it would have resulted in a conclusion that  $q \neq 1$ . From these reports, it is evident that statistical analysis and phenomenological modeling are insufficient to deduce the minimum requirements for HA-mediated fusion.

Here, we attempt to resolve these apparent contradictions through an algorithmic systems biology approach,<sup>14</sup> which mimics the logic of the biological system. In such an approach, biological objects and processes are transformed into objects and instructions in an executable program (Fig. 4). Such an approach is especially suited to the current case, where the knowledge of the steps involved is comparatively extensive. To setup the model, we use PABM, a formal language inspired by membrane processes.<sup>15</sup> PABM permits the representation of biological, membrane-bound objects as dynamic, nested compartments that can merge or split, and from which contents move in and out (Fig. 5A). Changes in compartment topology result from specific interactions of processes on the membranes of compartments; a biological example of such specific interaction is the interaction of a fusion peptide with its target, which precludes fusion (Fig. 5B). Given, however, that a PABM executor is currently under development, we mapped the model to PRISM reactions to check its behavior. Our model yields  $\omega = 6$ ,  $q = 3$ , where  $q$  is comprised of two free and one bound trimers. Apart from providing a possible resolution to the contradictions arising from data analysis, we were able to perform *in silico* experiments of previously untested scenarios, specifically, the effect of varying the surface density of SA. Our model yields a range of SA surface densities at which fusion can still occur. This might be able to explain the pathology of influenza in non-respiratory tract tissue and also be used as a criterion for determining if some individuals have a selective advantage against influenza. These results demonstrate the potential of algorithmic systems biology approaches in data interpretation and predictive modeling.

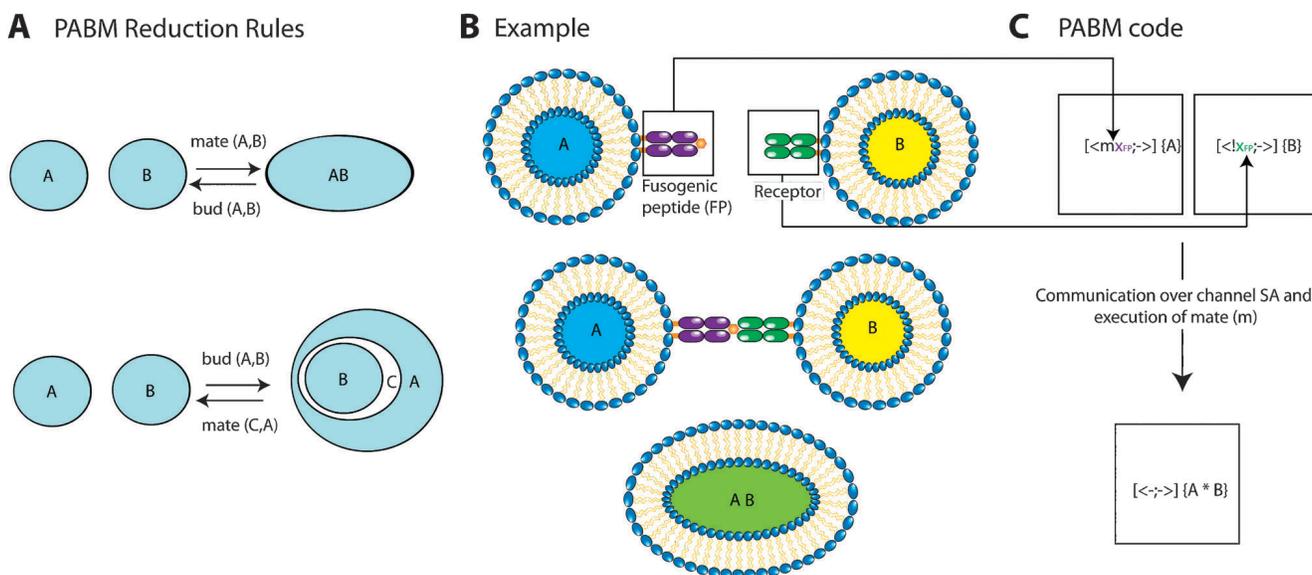
## 2 Methods

### 2.1 Computational modeling in PRISM

PRISM is a probabilistic and symbolic model checker which permits the analysis of all possible behaviors of the system.<sup>16</sup> Apart from the advantages provided by its model checking feature, it also has a simulation engine.<sup>17</sup> Given that the *in vitro* reactions we wish to model include one strictly membrane-related event, it is possible to map all events to biochemical-type reactions. We chose PRISM because it has the combined model checking and simulation features. The simulation feature permits us to quickly perform a sanity check of the system behavior, and to adjust initial parameter estimates. Model checking then allows us to explore all possible states and transitions, and allows us to determine if a certain property holds for a system.<sup>16</sup> Furthermore, it allows us to evaluate the effects of parameter changes on the probability of having fusion events. Finally, there have been numerous precedents for the use of PRISM in the modeling and analysis of biological pathways, including a previous mapping from another model involving compartments.<sup>18–20</sup>



**Fig. 4** Modeling workflow based on an algorithmic approach. Biological systems, which are described in terms of qualitative models ("cartoons"), as well as reaction stoichiometries and rates, are abstracted as objects, properties and algorithms that can be coded using a suitable language and executed. The behavior and reliability of the model can be evaluated through model checking and verification, respectively.



**Fig. 5** An overview of PABM. PABM is a formalism that addresses the need to intuitively express biological processes involving membranes. The basic operations of PABM on compartments, known as reduction rules, are fusion (mate) and fission (bud) (A). These rules are implemented in response to specific communications between actions on membranes. Actions define which compartments can interact, as well as the fusion and fission capabilities of the membranes they are associated with. (B) In a simple biological example, a fusogenic peptide FP on the surface of a biological system A may be represented as an action with an instruction for mate (designated  $m_{x_{FP}}$  in PABM code). Following its interaction with another system B that has the appropriate receptor for FP (designated  $!x_{FP}$  in PABM code), mate is executed, and both the membranes and contents of A and B mix. Note the one-to-one correspondence between the biological system and PABM code, where the objects corresponding to A and B are likewise combined (C).

## 2.2 Biochemical processes

The biological processes associated with the HA-mediated fusion setups are shown in Fig. 1, where each of the reactions described below are indicated. These reactions were initially formulated in the PABM formalism, then mapped into the guarded commands required in PRISM; the properties analyzed, namely the probability of having  $Virus_{FP}$  and  $Virus_{LC}$  at time  $t$  for virus–cell and cell–cell fusion setups were expressed

in continuous stochastic logic (CSL)<sup>16</sup>

(1) Virus binding and unbinding

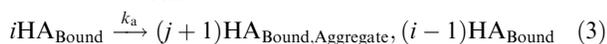


$HA_{Bound} + SA_{Bound}$  are counted as  $HA_{Bound} \cdot SA_{Bound}$

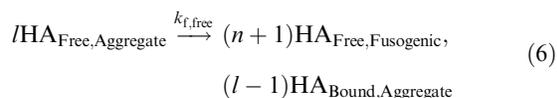
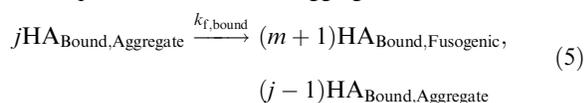


(2) pH-induced HA aggregation and conformational change

If  $i \geq 1$ , where  $i$  is the number of  $\text{HA}_{\text{Bound}}$  and  $k$  is the number of  $\text{HA}_{\text{Free}}$ :



where  $j$  and  $l$  are the number of bound and free HA trimers in an aggregate, respectively, and where  $j$  and  $l = 0$  at the start of the simulation. Both aggregation reactions are preceded by a synchronization guard of rate 1.0 to ensure that all anchor points between the fusing membranes have been established prior to HA clustering. If  $\text{HA}_{\text{Bound,Aggregate}} \geq \min \text{HA}_{\text{Bound,Aggregate}}$  and  $\text{HA}_{\text{Free,Aggregate}} \geq \min \text{HA}_{\text{Free,Aggregate}}$  and  $\text{HA}_{\text{Bound,Aggregate}} + \text{HA}_{\text{Free,Aggregate}} \geq \omega$ , where  $\min \text{HA}_{\text{Bound,Aggregate}}$  and  $\min \text{HA}_{\text{Free,Aggregate}}$  are user-defined and  $\omega$  is equal to the minimum aggregate size:



where  $m$  and  $n = 0$  at the start of the simulation. Note that there is a concurrent addition to the number of trimers in an aggregate and a subtraction from the corresponding pool of trimers that were previously not associated with any aggregate. This is denoted by the comma on the right-hand side of the equation.

(3) Fusion pore (FP), lipid channel (LC) and fusion site (FS) formation

If  $\text{HA}_{\text{Bound,Fusogenic}} \geq \min \text{HA}_{\text{Bound,Fusogenic}}$  and  $\text{HA}_{\text{Free,Fusogenic}} \geq \min \text{HA}_{\text{Free,Fusogenic}}$  and  $\text{HA}_{\text{Bound,Fusogenic}} + \text{HA}_{\text{Free,Fusogenic}} \geq q$ , where  $\min \text{HA}_{\text{Bound,Fusogenic}}$  and  $\min \text{HA}_{\text{Free,Fusogenic}}$  are user-defined and  $q$  is a subset of  $\omega$  that undergoes a conformational change:



where  $\text{Virus}_{\text{FP}}$ ,  $\text{Virus}_{\text{LC}}$  and  $\text{Virus}_{\text{FS}}$  represent virus particles containing a fusion pore, a lipid channel and a fusion site, respectively.

## 2.3 Model assumptions and parameter estimates

**2.3.1 Inclusion of binding step.** Fusion experiments involve pre-binding of viruses or HA-expressing cells to the target membrane, making the virus binding reactions appear unnecessary. However, pre-binding does not prevent additional binding events from taking place in the gap between the pre-binding step and the pH drop.<sup>21</sup> Furthermore, the explicit representation of bound and unbound HA trimers is necessary for determining the subset of bound HA trimers in  $\omega$  and  $q$ .

**2.3.2 HA and SA surface densities.** The estimate of the number of HA trimers/virion was taken from independent

reports by Imai *et al.*, Saitakis and Gizeli and Taylor *et al.*<sup>10,22,23</sup> Other parameters, such as the HA and SA surface density at the contact area (Table 1) were obtained from information in the original papers, as well as estimates in a previous modeling paper.<sup>24</sup> We first performed simulations using these values; we then took the final HA : SA values obtained for successful fusions within the expected time scale and used this ratio in model checking. Given computing constraints in the model checker of PRISM, where values of the order of a hundred molecules for this model result in an out-of-memory error, we scaled down both the HA and SA values to the order of 15 and 5, respectively, reflecting the average 74 HA : 30 SA ratio that results in successful fusion.

**2.3.3 Initial parameter estimates.** Most rates for each of these transitions, with the exception of  $k_b$  and  $k_a$ , are either not available in the literature, or could not be estimated from literature values (Table 2). Initial parameter estimates were derived from known rates of diffusion,<sup>8</sup> which presumably affects the aggregation rate,  $k_a$ , as well as predicted rates of binding,  $k_b$ .<sup>25</sup> Initial values for the acid-induced conformational change,  $k_f$ , and fusion pore formation,  $k_{fp}$ , were based on parameters obtained from fits reported by Bentz.<sup>5</sup> In the case of  $k_f$ , we make a distinction between  $k_{f,\text{bound}}$  and  $k_{f,\text{free}}$  to allow us to test the cases  $k_{f,\text{bound}} \ll k_{f,\text{free}}$  and  $k_{f,\text{bound}} = 0$ , given that there is no conclusive experimental evidence regarding the ability or inability of bound HA molecules to undergo a conformational change.<sup>26,27</sup> Nonetheless, if it is able to undergo the conformational change, it could be reasonably expected to be slower.<sup>26</sup> Consequently, we have assumed that it has a rate 1/100 of the original  $k_f$  value. For simplicity, only forward reactions were considered, although reactions (1)–(7) are known to be reversible.

Since PRISM does not have a built-in parameter optimization toolkit, derived parameters were obtained using different combinations of parameter ranges; these ranges were chosen based on preliminary runs evaluating the model behavior when a single parameter is varied, while the others are held constant

**Table 2** Model parameters and rates

| Reaction              | Parameters  | Initial rate/s <sup>-1</sup> | Fitted rate/s <sup>-1</sup>                 |
|-----------------------|---|------------------------------|---|
| Binding               | 74 HA, 223 SA for virus–cell fusion experiments; 15 HA, 5 SA for model-checking 18 HA, 446 SA to 30 HA, 446 SA for cell–cell fusion experiments; 15 HA, 5 SA for model-checking | 0.2 <sup>25</sup>            | 34.81                                       |
| Unbinding             | n/a   | n/a                          | 0.0001–0.25                                 |
| HA aggregation        | $\omega$ from 5 to 9  | 830 <sup>8,25</sup>          | 100, virus–cell, 0.00225–0.00765, cell–cell |
| Conformational change | $q$ from 1 to 3   | —                            | 6.25  |
| Transition to FP      | n/a   | —                            | 0.9025                                      |
| Transition to LC      | n/a   | n/a                          | 1.1025                                      |
| Transition to FS      | n/a   | n/a                          | n/a   |

(data not shown). The data reported by Imai *et al.* were initially fitted; the set of parameters associated with the best fits, with the exception of the aggregation (presumed slower), were then used in fitting the data reported by Melikyan *et al.*

A summary of model parameters and model-derived rates is presented in Table 2.

## 2.4 Reducing the solution space

To determine  $\omega$ ,  $q$ , and the individual states (bound or unbound) of each trimer within  $\omega$  and  $q$ , the user-defined parameters  $\min \text{HA}_{\text{Bound,Aggregate}}$ ,  $\min \text{HA}_{\text{Free,Aggregate}}$ ,  $\min \text{HA}_{\text{Bound,Fusogenic}}$  and  $\min \text{HA}_{\text{Free,Fusogenic}}$  can be varied to reflect all possible cases. For  $\omega$ , we initially tested the values ranging from 5 to 9 (*viz.*  $6 \pm 1$  and  $8 \pm 1$ ). It is assumed that an aggregate can be comprised of both bound and free HA particles. We also assumed that  $q$  can include both bound and free HA, though  $k_{f,\text{free}}$  is significantly faster than  $k_{f,\text{bound}}$ . Taken together, and eliminating cases that are not biologically plausible (*viz.* cases where none of the HA molecules in  $\omega$  are bound) the different combinations result in a total of roughly 235 test cases (39 possible combinations for  $\min \text{HA}_{\text{Bound,Aggregate}} + \min \text{HA}_{\text{Free,Aggregate}}$  yielding a value from 5 to 9, each considered in the context of an average of 6 possible cases of  $\min \text{HA}_{\text{Bound,Fusogenic}} + \min \text{HA}_{\text{Free,Fusogenic}}$  for the range from 1–3). Finally, to eliminate even more unlikely scenarios, we used the fastest reaction for each test set; for  $\omega = 8$ , for example, the fastest reaction occurs when  $\min \text{HA}_{\text{Free,Aggregate}} = 7$ , and  $q = 1$ , where the only trimer undergoing a conformational change is free. Using this strategy, we initially determined the most probable value of  $\omega$ , then used these values for determining  $q$ .

## 2.5 Parameter sensitivity analysis

Sensitivity analysis was performed to estimate the reliability of the model predictions. Here, we used local sensitivity analysis adapted for the stochastic case. Briefly, parameter values were changed one at a time, while keeping the rest fixed. Sensitivity indices  $S_a$ , which represent the sensitivity of the output to a change in each parameter  $P_i$ , were calculated based on the standard:<sup>28</sup>

$$S_a = \frac{\partial Y}{\partial P_i} \quad (9)$$

where  $\partial Y$  is calculated as the changes in the output resulting from the substitution of the reference parameter with new parameters in incremental ratios, and  $\partial P_i$  is the difference between the reference and the new parameter. We compared the full distributions of outputs obtained from model checking for each of the parameters in order to account for changes in the shape of the distribution.

## 2.6 Variable SA experiments

For the variable SA experiment, we changed the values of the effective HA : SA surface density at the contact area from an original estimated value of 3.0, to values between 0.75 and 15.0, while holding the values of  $\omega$  and  $q$ , obtained using the procedure in Section 2.4, constant.

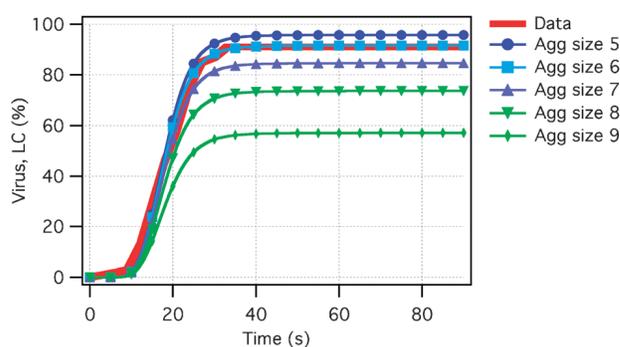
## 3 Results

### 3.1 A minimum aggregate size of six trimers is required for the fusion pore

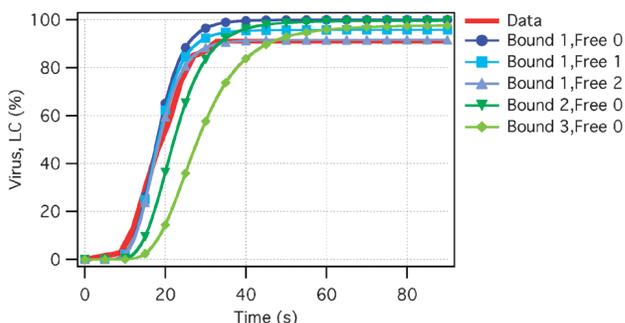
To determine  $\omega$ , we initially took the range of 5–9 trimers as possible minima required for the transitions in eqn (5) and (6) to occur. As shown in Fig. 6 a requirement for  $\omega \geq 8$  and above would not be able to account for the observed fusion kinetics, although this does not mean that aggregates of this size would not result in fusion. Rather, it simply indicates that majority of the fusion events would have to involve complexes of a smaller size. Presumably, it would require more time to assemble an aggregate of this size. In contrast,  $\omega = 6$  closely fits the data. Furthermore, this size is consistent with electron microscopy-based approximations of the pore size formed by the so-called HA rosettes, which can be generated by solubilizing the viral membrane with detergent then sparsely redistributing them across synthetic liposomes.<sup>29</sup>

### 3.2 A minimum of three trimers in the fusion pore have to undergo a conformational change to become fusogenic

Using the results described in Section 3.1, we simultaneously varied the values of  $\min \text{HA}_{\text{Bound,Fusogenic}}$  and  $\min \text{HA}_{\text{Free,Fusogenic}}$  to determine  $q$  (Fig. 7). A value of  $q = 3$ , comprised of one bound and two free HA trimers, fits the data. Interestingly, different values of  $q$  do not significantly affect the level of



**Fig. 6** Simulation of fusion data reported by Imai and Floyd (red) with the assumption that  $\omega$  is 6 or 8, with at least one bound HA trimer in each case. In the case of  $\omega = 8$ , fusion is still observed, but does not fit the data.



**Fig. 7** Dependence of fusion kinetics on  $q$  for  $\omega = 6$ . Of all the possible combinations, an aggregate comprised of at least one bound and five free HA trimers, of which three (one bound, two free) must undergo a conformational change, best describes the data.

fusion as much as  $\omega$ , but instead causes a shift in the time at which saturation is reached.

### 3.3 Model validation using cell–cell fusion experiments

We next tried to verify our results by using the predictions in Sections 3.1 and 3.2 to fit the data reported by Melikyan *et al.*,<sup>12</sup> which were obtained using cell–cell fusion measurements, where HA-expressing cells are used instead of viruses. Cell–cell fusion experiments differ from virus–cell fusion experiments in terms of the HA and SA surface densities at the contact area. These are also characterized by slower kinetics because of the lower HA surface density, as well as the presence of other proteins that can influence the fusion kinetics.<sup>8,11</sup> Nonetheless, all mechanisms starting from the point where the aggregate is assembled (3 and 4, Fig. 1) are identical.<sup>5</sup> It should thus be possible to capture the behavior of both HA-mediated virus–cell and cell–cell fusion using a single model. For this, we varied  $k_a$ , which is presumably slower. However, due to the memory constraints in PRISM, we had to scale down values of HA and SA, such that the effective HA : SA surface density ratio is maintained (Table 2), instead of using the actual values indicated in Table 1. The best fits for both data sets are still  $\omega = 6$ , with  $k_a$  values ranging from 0.00225–0.00765 s<sup>-1</sup> (Fig. 8). Other cell–cell fusion measurement data<sup>8,13</sup> were presumed transformable to allow direct comparison with the data reported by Melikyan *et al.*, and were no longer modeled in this paper.<sup>7</sup>

### 3.4 Parameter sensitivity analysis

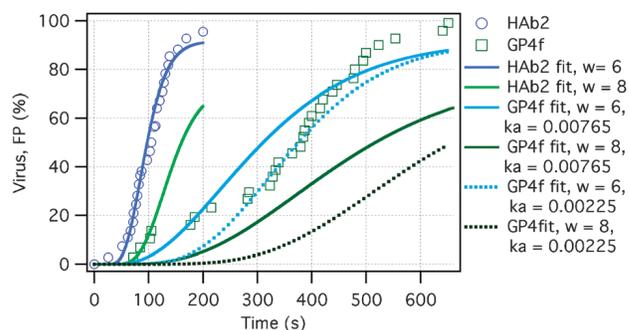
Following model validation, we performed sensitivity analysis on estimated parameters. The highest sensitivity index was obtained for  $k_f$ , which is consistent with the presumed role of HA conformational change as a rate-limiting step.<sup>5</sup> Changing the other rates without changing any of the non-estimated parameters, such as HA and SA surface density, does not materially affect the results (Fig. 9). The relatively low parameter sensitivities are indicative of the robustness of the model.

### 3.5 Effect of SA surface density on fusion kinetics

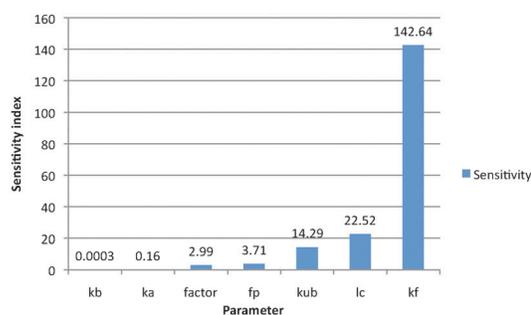
Another application that we found for the model is to check the effect of SA surface density on influenza fusion. For this, we used the parameters obtained for the virus–cell fusion setup and assumed  $\omega = 6$  and  $q = 3$  (1 bound, 2 free), while varying the HA : SA surface density between 0.75 to 15.0.\*\* Of these concentrations, only HA : SA ratios between 5.0 and 2.5 resulted in at least 90% fusion (Fig. 10). At HA : SA concentrations lower than 1.67, fusion decreases dramatically, with almost no fusion occurring at HA : SA = 1.25 and below. This decrease in fusion efficiency is a necessary consequence of the predicted requirement for at least two free trimers in  $q$ ; with more SA molecules available, the incidence of HA binding would be higher, and it would presumably require more time for the fusogenic complex to be assembled,

|| This was not assumed *a priori* in our model.

\*\* Experiment involving virus–cell fusion experiments have an average HA : SA surface density ratio of 3.0.



**Fig. 8** Simulation of the Melikyan fusion data using parameters obtained from the Imai data, with the exception of HA and SA surface densities, and the  $k_a$  value, which was presumed to be slower than in virus–cell fusion experiments.  $k_a$  values between 0.00225–0.00765 s<sup>-1</sup>, which are approximately of order 10<sup>4</sup> slower than  $k_a$  values for virus–cell fusion data, were obtained. The results of the model with  $\omega = 8$ ,  $q = 2$ , as reported by Bentz,<sup>5</sup> are shown for comparison.

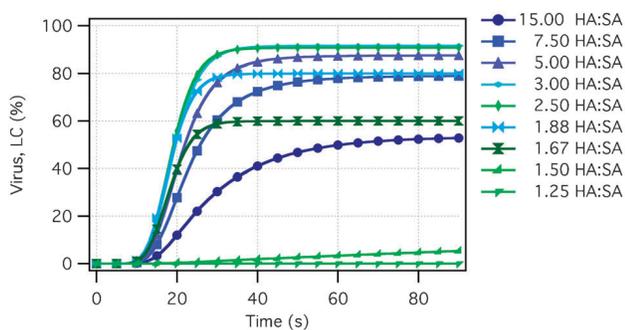


**Fig. 9** Sensitivity indices of estimated parameters. The model is sensitive to changes in  $k_f$ , the rate of HA conformational change. Parameters tested were rates of binding ( $k_b$ ), unbinding, ( $k_{ub}$ ), aggregation ( $k_a$ ), fusion pore formation ( $k_{fp}$ ) and lipid channel formation ( $k_{lc}$ ), as well as the factor by which  $k_f$  is decreased when bound HA undergo a conformational change.

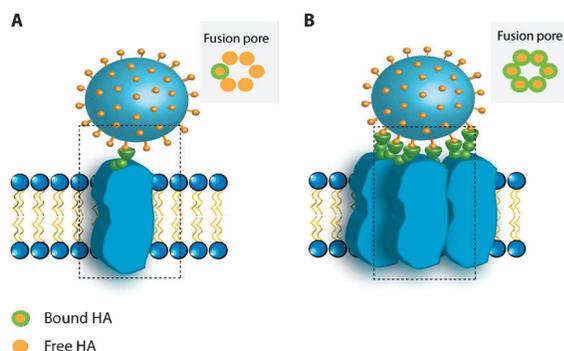
if this has not yet been physically prevented by bound molecules at the contact site (Fig. 11).

These results are partly contrary to those obtained by Schreiber *et al.*, who predicted that a higher surface density of SA (*viz.* receptor density) is slightly more efficient than increasing the HA concentration in accelerating the fusion process.<sup>25</sup> It is true that for HA : SA ratios between 15.0 and 5.0, the fusion process is accelerated, and that the extent of fusion increases. However, at HA : SA ratios lower than 2.5, the effect is reversed. A probable reason for this discrepancy is their omission of the HA conformational change requirement for fusion. Furthermore, our results are consistent with earlier studies that have observed lower incidences of fusion when fusion partners with an extremely high SA content were used.<sup>30</sup>

Finally, the predictions may be significant with respect to the pathology of infection of certain types of influenza, which are not limited to tissue in the respiratory tract, but have also been observed in cells in the brain, lymph nodes, liver, kidney, spleen and intestine, which express SA receptors in an appreciable number. However, there was no infection in the esophagus, heart and bone marrow, even if both esophageal and cardiac tissues are in closer proximity to the respiratory tract than the kidney,<sup>4</sup> presumably due to the unavailability or insufficiency of



**Fig. 10** SA surface density affects viral fusion efficiency. For the variable SA experiments,  $\omega = 6$  and  $q = 3$  were kept constant. The HA : SA surface density ratios were then varied from 15.0 down to 0.75. Only values between 5.0 and 2.5 HA : SA resulted in at least 90% fusion; between 1.50 and 0.75 HA : SA, no appreciable fusion is expected to occur, given that most of the HA molecules would be bound, and the requirement for two free HA trimers in  $q$  is unlikely to be met.



**Fig. 11** Influence of SA surface density on fusion-permissive (A) and non-permissive (B) pore formation. Dotted lines represent the contact area in which the fusion pore is formed. The insets show the pores that are created when the SA surface density is lower (A); a higher surface density of SA leads to more bound HA trimers, making it difficult for the requirement of at least two free HA molecules to occur within a complex to be fulfilled.

SA receptors. It is also possible that the variation of SA surface density among individuals confer selective advantages against influenza. For instance, RBCs of thalassemia patients and diabetes patients have been reported to have lower sialic acid content.<sup>31,32</sup> Given that most fusion experiments were performed using red blood cell (RBC) ghosts, the predicted effect may be tested by performing fusion studies using RBC ghosts from patients with these diseases; artificial model membranes containing varying concentrations of purified glycoprotein, the main sialylated protein of the RBC, could also be used for verifying our predictions.

## 4 Discussion

Determining the minimum requirement for a virus to create a fusion pore would provide important insights into the first line of influenza pathogenesis. Knowledge of this minimum requirement would also have interesting applications in drug and artificial gene delivery vector design, where endosomal escape has remained a perennial problem.<sup>33</sup> Several groups

have worked on determining values of  $\omega$  and  $q$  for the past 20 years, starting with measurements on HA-trimer expressing cells, and later, on virus-like or virus particles, once an adequate visualization technology was available. However, the experimental setups designed for this purpose were very different, and have all been documented to have an effect on fusion kinetics.<sup>7,34,35</sup> Furthermore, the statistical models and phenomenological methods used in data analysis were also widely varied. It is consequently not surprising that the data obtained appear to be very different at a first glance and that the conclusions derived from them appear contradictory.

A possible solution would be to create a model of the processes, for which consistent parameters could be obtained for at least one virus–cell and one cell–cell fusion experiment. Previous efforts to model the process used mass action kinetics to describe the fusion intermediates starting from the conformational change within the HA aggregate that leads to FP formation.<sup>5,24</sup> The formation of the HA aggregate is not included in the model as a step explicitly, since they have assumed that this is not a rate-limiting step.  $\omega$  is instead estimated using a nucleation model. The first of the two models<sup>5</sup> yielded a value of  $\omega = 8$ , and a value of  $q = 2$  or 3. A succeeding paper<sup>24</sup> that builds on this model by analyzing additional cell–cell fusion experiments yields  $q = 2$ , with the assumption that  $\omega = 8$ .

Here, we try a different approach where we create a stochastic model that includes aggregate formation explicitly. The inclusion of the aggregation step is necessary if we want to derive a parameter set for a model that can fit both virus–cell and cell–cell fusion data. Apart from the fact that the HA surface densities in viruses and cells are different, only the aggregate formation rate,  $k_a$ , no matter how fast it is compared to the rate of conformational change,  $k_f$ , is the only other thing that can vary between the two setups. All the other steps, from the formation of  $q$  to FS, should be the same. In fact, it is noted in an earlier paper of Bentz<sup>36</sup> that the aggregation step in HA-expressing cells appears to be an unfavorable, probably highly reversible event. This is in stark contrast to the step in viruses, where HA trimers might even be almost pre-aggregated.<sup>36</sup> Furthermore, it would be necessary to know the states of HA in the aggregate (*viz.* bound or unbound) if we want to know which of these participate in the formation of  $q$  fusogenic units.

Our model is sensitive to both  $\omega$  and  $q$ , with the extent of fusion being dependent on the aggregation step. On one hand, this dependence on  $\omega$  even for virus–cell fusion setups can appear counter-intuitive, since the density and relative proximity of the trimers on a virus surface could make them practically pre-aggregated. However, if one thinks of it as a reaction at least in 2D, then it could be that constructing a fusogenic aggregate might be slightly more complicated based on how many bound molecules are there at the contact area to begin with. In such a case, the dependence of the kinetics on  $\omega$  could be explained. The extension of this model to a lattice, as reported by Schreiber *et al.*<sup>25</sup> would be particularly useful in tackling such a question.††

†† This article is not discussed in detail since its purpose is not so much as to determine the smallest fusogenic unit as it is to present a new technique for approaching the problem.

There are discrepancies between the rates obtained using fits from this model and those in that reported by Bentz,<sup>5</sup> which could probably be naturally expected from the fact that the model structure and the assumptions held are different. The fit generated by Bentz for the GP4f data is, of course, superior to the fit that we have obtained, but this might have been partly due either to the overly scaled-down approximation of the number of HA and SA molecules, or to an overfitting of potentially noisy data. Nonetheless, the ability of the model to closely capture both virus–cell and the general behavior of the cell–cell fusion experiments, while keeping the parameters that are expected to be constant, is promising. In the future, a more complete comparison of the two models, towards which modeling the virus–cell fusion experiments using the methods of Bentz would be a first step, would be particularly interesting. It would likewise be interesting to factor in the involvement of HA trimers outside the fusion site in fusion pore expansion.<sup>37</sup> Finally, we are working on creating an experimental setup to verify either of the predictions. In the advent of technologies that permit the manipulation of individual molecules with nanometre precision, it would not be so remote to conceptualize a nanoparticle with a defined number of hemagglutinin trimers at its surface. Coupled with microscopy that allows the tracking of individual HA trimers,<sup>38</sup> such a technique should be able to settle the question of the minimal fusion requirements definitively, while functioning as a litmus test for the significance of the results obtained from modeling processes of this scale.

## Acknowledgements

MPD thanks the Deutscher Akademischer Austausch Dienst for her PhD scholarship. AD is supported by the Swiss National Foundation of Science grant number 31003A-125457. The authors wish to thank Masaki Imai and Daniel Floyd for graciously providing their original data, and Thomas Ligon for his valuable input in validating the model and the manuscript.

## References

- 1 T. Söllner, *Curr. Opin. Cell Biol.*, 2004, **16**, 429–435.
- 2 M. Barocchi, V. Masignani and R. Rappuoli, *Nat. Rev. Microbiol.*, 2005, **3**, 349–358.
- 3 J. J. Skehel and D. C. Wiley, *Annu. Rev. Biochem.*, 2000, **69**, 531–569.
- 4 Y. Piwpankaew, Y. Monteerarat, O. Suptawiwat, P. Puthavathana, M. Uipresertkul and P. Auewarakul, *APMIS*, 2010, **118**, 895–902.
- 5 J. Bentz, *Biophys. J.*, 2000, **78**, 227–245.
- 6 Y. Krishnamachari, S. M. Geary, C. D. Lemke and A. K. Salem, *Pharm. Res.*, 2011, **28**, 215–236.

- 7 A. Mittal, E. Leikina, J. Bentz and L. Chernomordik, *Anal. Biochem.*, 2002, **303**, 145–152.
- 8 T. Danielli, S. Pelletier, Y. Henis and J. White, *J. Cell Biol.*, 1996, **133**, 559.
- 9 D. Floyd, J. Ragains and J. Skehel, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 15382–15387.
- 10 M. Imai, T. Mizuno and K. Kawasaki, *J. Biol. Chem.*, 2006, **281**, 12729–12735.
- 11 S. Gunther-Ausborn, P. Schoen and I. Bartoldus, *J. Virol.*, 2000, **74**, 2714–2720.
- 12 G. B. Melikyan, W. D. Niles and F. S. Cohen, *J. Gen. Physiol.*, 1995, **106**, 783–802.
- 13 R. Blumenthal, D. Sarkar and S. Durell, *J. Cell Biol.*, 1996, **135**, 63–71.
- 14 J. Fisher and T. Henzinger, *Nat. Biotechnol.*, 2007, **25**, 1239–1249.
- 15 M. David, J. Bantang and E. Mendoza, *Trans. Comput. Syst. Biol.*, 2009, **11**, 164–186.
- 16 M. Kwiatkowska, G. Norman and D. Parker, *Symbolic Systems Biology, Probabilistic Model Checking for Systems Biology*, Jones and Bartlett, 2010, pp. 31–59.
- 17 T. Pronk, E. de Vink, D. Bosnacki and T. Breit, *Proc. MTCoord 2007*, Paphos, 2007.
- 18 M. Kwiatkowska, G. Norman and D. Parker, *ACM SIGMETRICS Performance Evaluation Review*, 2008, vol. 35, pp. 14–21.
- 19 F. Romero-Campero, M. Gheorghe, L. Bianco, D. Pescini, M. Pérez-Jiménez and R. Ceterchi, *Membrane Computing*, Springer, Berlin/Heidelberg, 2006, vol. 4361, pp. 477–495.
- 20 J. Krivine, R. Milner and A. Troina, *Proc. of MFPS'08, 24th Conference on the Mathematical Foundations of Programming Semantics*, 2008, pp. 73–96.
- 21 M.-C. Giocondi, F. Ronzon, M. Nicolai, P. Dosset, P.-E. Milhiet, M. Chevalier and C. L. Grimmelc, *J. Gen. Virol.*, 2010, **91**, 329.
- 22 H. Taylor, S. Armstrong and N. Dimmock, *Virology*, 1987, **159**, 288–298.
- 23 M. Saitakis and E. Gizeli, *Eur. Biophys. J.*, 2010, **40**, 209–215.
- 24 A. Mittal and J. Bentz, *Biophys. J.*, 2001, **81**, 1521–1535.
- 25 S. Schreiber, K. Ludwig, A. Herrmann and H. Holzhütter, *Biophys. J.*, 2001, **81**, 1360–1372.
- 26 H. Ellens, J. Bentz, D. Mason, F. Zhang and J. White, *Biochemistry*, 1990, **29**, 9697–9707.
- 27 B. Millar, L. Calder, J. Skehel and D. Wiley, *Virology*, 1999, **257**, 415–423.
- 28 A. B. Massada and Y. Carmel, *Ecol. Modell.*, 2008, **213**, 463–467.
- 29 R. Ruigrok, N. Wrigley, L. Calder, S. Cusack, S. Wharton, E. Brown and J. Skehel, *EMBO J.*, 1986, **5**, 41.
- 30 D. Alford, H. Ellens and J. Bentz, *Biochemistry*, 1994, **33**, 1977–1987.
- 31 I. Kahane, E. Ben-Chetrit, A. Shifter and E. Rachmilewitz, *Biochim. Biophys. Acta, Biomembr.*, 1980, **596**, 10–17.
- 32 K. Z. Boudjeltia, M. Piagnerelli, P. Piro, D. Bastin, E. Carlier, P. Lejeune and M. Vanhaeverbeek, *Crit. Care*, 2000, **4**, 17.
- 33 K. Douglas, *Biotechnol. Prog.*, 2008, **24**, 871–883.
- 34 A. Puri, F. Booy, R. Doms, J. White and R. Blumenthal, *J. Virol.*, 1990, **64**, 3824.
- 35 V. I. Razinkov, G. B. Melikyan and F. S. Cohen, *Biophys. J.*, 2008, **77**, 3144–3151.
- 36 J. Bentz, *Biophys. J.*, 1992, **63**, 448–459.
- 37 E. Leikina, A. Mittal, M. Cho, K. Melikov, M. Kozlov and L. Chernomordik, *J. Biol. Chem.*, 2004, **279**, 26526.
- 38 S. Hess, T. Gould, M. Gudheti, S. Maas, K. Mills and J. Zimmerberg, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 17370.

# Appendix F

## Manuscript M1

Maria Pamela Dobay, Alexandra Schmidh, Eduardo Mendoza, Thomas Bein, Joachim O. Rädler

**Cell-type dependence of protoporphyrin IX-induced endosomal escape kinetics**



# Cell type dependence of protoporphyrin IX-induced endosomal escape kinetics

Maria Pamela Dobay<sup>1\*</sup>, Alexandra Schmidt<sup>2\*</sup>, Eduardo Mendoza<sup>1</sup>, Thomas Bein<sup>1</sup>, Joachim O. Rädler<sup>1</sup>

<sup>1</sup>Faculty of Physics and Center for NanoScience, Ludwig Maximilians University, Geschwister-Scholl-Platz 1, 80539, Munich, Germany

<sup>2</sup>Department of Chemistry and Center for NanoScience, Ludwig Maximilians University, Butenandstrasse 5-13, 81377 Munich, Germany

\*These authors contributed equally to this work

## Abstract

We investigated uptake and individual endosome lysis events in fibroblast, normal and carcinoma cell lines using a colloidal mesoporous silica (CMS) nanoparticle (NP)-based reporter system. Endosome lysis was induced through the activation of protoporphyrin IX (PPIX). Surprisingly, this release-on-demand system resulted in more broadly-distributed lysis times than expected, particularly for Renca, a renal carcinoma cell line. An analysis of the NP load per endosome, endosome size and uptake characteristics indicate that Renca cells not only take up a lower amount of NPs in comparison with the fibroblast cells, but also have larger endosomes, and a lower NP load per endosome. We then extended an existing stochastic pi calculus model of gold NP intracellular distribution to understand how much factors that cannot be directly measured, such as variations in the PPIX load per NP, affect the distributions. Model results indicate that the lysis time distribution is primarily determined by the minimum net PPIX required to burst an endosome, a factor influenced by the NP load per endosome, as well as the endosome size.

## Keywords

Supported lipid bilayer, fluorimetric reporter system, colloidal mesoporous silica nanoparticles, endosomal escape, escape-on-demand systems, rule-based models, sensitivity analysis

## Introduction

Endosomal escape has been identified as one of the main bottlenecks in gene and drug delivery [1, 2]. To date, however, there are few papers that investigate single endosome lysis in real time because of low signal-to-noise ratios [3]. Previous studies have recorded the intracellular trafficking of individual nanoparticles and various delivery vectors [4-6], measured endosome or cellular contents *ex vivo* [2, 7], or tracked individual viruses as they move across the cell [8]. In these protocols, individual components of the virus or the vector, as well as cellular components, are labeled [8, 9]. In all studies involving delivery vector tracking [4-6], confocal laser scanning microscopy was used, since dye localization on a single plane is insufficient to definitively establish particle internalization [10]; it was reported in Akita et al. that at least 20 z-slice images were required for their method, and that at least 30 cells were needed to be imaged in this manner to obtain statistically meaningful results [4]. If this method were translated into a 30-hour observation period, as in [11], this would entail the capture, consolidation and analysis of 216,000 images.

Due to such limitations, most of the studies have concentrated on the spatial distribution of the material [5, 6], or on the trafficking mechanisms used instead [4, 8]. Endosome escape efficiency is inferred from a minimal amount of sampling points [4]. Uptake events are

likewise rarely recorded in real time, again due to the low signal-to-noise ratio; real time recording frequently requires confocal or total internal reflection microscopy [8]. It is more typical for endosomal escape rates to be estimated from data fits [2]. Combining direct analysis of uptake and endosomal escape will not only yield time distributions of the escape events, but also information on the vector load per endosome, and the influence of endosome size on escape. Availability of such information for vectors tested on different cell lines would be important in evaluating their performance.

To obtain real-time information on individual lysis events, we needed a reporter system that could be co-delivered with, or itself used as a delivery vector. An ideal detector system should have distinct signals for each target compartment, and the signals produced should not immediately diffuse into the cytosolic milieu. For this, we chose to modify a recently designed colloidal mesoporous silica (CMS) nanoparticle (NP) system to create a fluorogenic detector.

CMS NPs have attracted great attention in the past years as potential drug delivery vehicles due to the non-toxicity and biocompatibility of silica [12-14]. Our spherical CMS NPs possess a controllable and well-defined uniform porosity, offer large surface areas, pore volumes, multifunctionality, and are small in size, a critical feature for endosomal uptake [15-17]. The surface of these NPs can be modified by inorganic and organic functionalization to give multifunctional hybrid materials [18-21]. Functionalization of the external surface of CMS with polymers such as polyethylene glycol (PEG) [22] and polyethyleneimine (PEI) [23], which increase colloidal stability by preventing non-specific protein adsorption [24-27], is an important factor for functional drug delivery vehicles. The release of various types of cargo from mesoporous silica materials ranging from plant-derived model-drugs [28] to poorly water-soluble cancer therapeutics like doxorubicin [29] and paclitaxel [30] has been studied intensively.

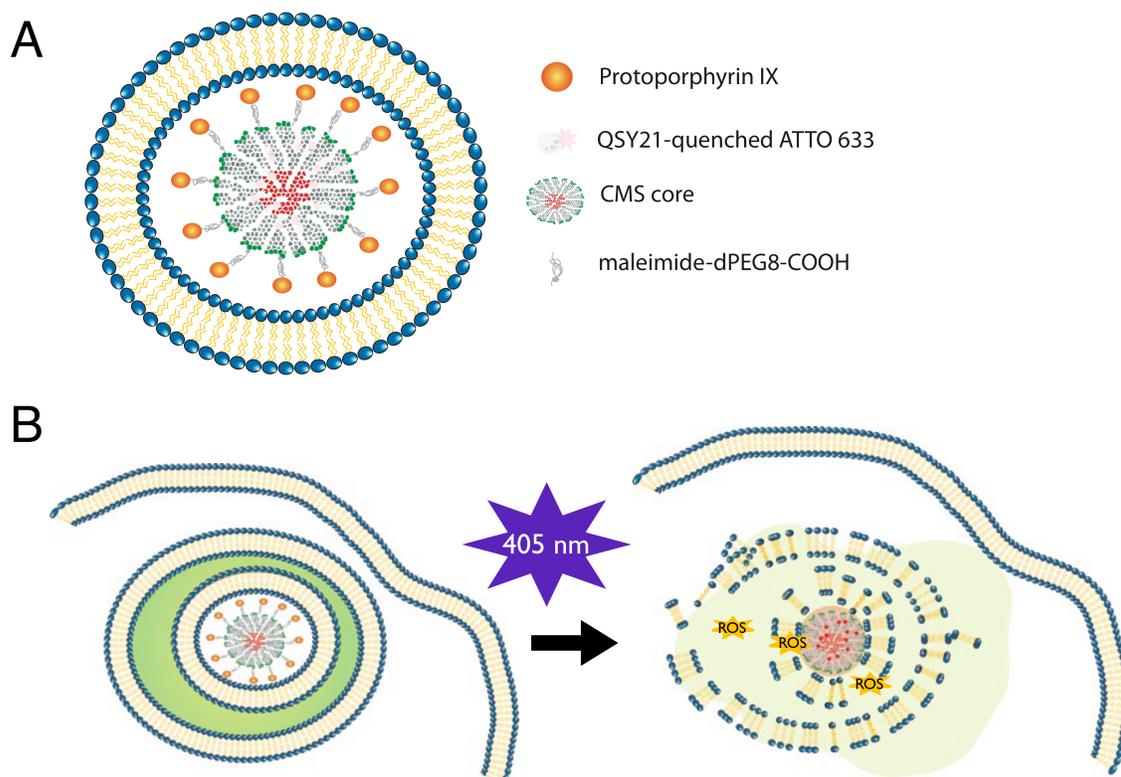
Cap systems on mesoporous silica acting as stimuli-responsive release mechanisms for controlled release have been widely studied in the past years [31, 32]. Inorganic nanoparticles [33], large molecules (i.e. cyclodextrins and rotaxanes) [34-36] and also polymers [37] have been used to prevent cargo from leaving the pore system. The opening-stimulus can be a change in pH [38], UV irradiation [39], or redox-activity [33]. A biomolecule-based enzyme-responsive cap system for mesoporous silica is based on the formation of the biotin-avidin complex at the pore openings of the host [40]. This cap system was extended further to a temperature-responsive release-on-demand system based on DNA strands [41]. It was also possible to successfully enclose the microtubule depolymerizing plant toxin colchicine in the mesopores of CMS by capping these with a supported lipid bilayer [28]. This approach employs the so-called solvent exchange induced self-assembly of lipids around the CMS resulting in a highly biocompatible silica-lipid hybrid system (SLB@CMS) [42]. NP-supported lipid bilayers exhibit a narrower size distribution and enhanced stability compared to liposomes [43-45].

Combining SLB@CMS with photochemical internalization (PCI) [46, 47] and photosensitizers (PS) leads to a highly potent drug delivery system [36]. PS generate reactive oxygen species (ROS) upon light-activation, leading to membrane rupture [48]. Recently, multifunctional CMS particles have been equipped with an on-board photosensitizer, namely amino-modified protoporphyrin IX (PpIX-NH<sub>2</sub>). This system has already been tested in multiple applications involving three cell active model drug systems.

Another possibility to prevent adsorbed compounds from leaching before successful delivery is the covalent attachment of cargo into the pores of mesoporous silica [29, 49]. The release of

the bee venom melittin attached to SBA-15 via a pH-responsive acetal-linker [50], and a redox-sensitive disulfide-linker [51] have been reported in the context of drug delivery. Many therapeutic peptides have thiol functionalities and nucleic acids can be modified with thiol moieties. These can be used to build disulfide-linkers to prevent adsorbed compounds from leaching. The reductive milieu in the endosome, redox-driven intracellular disulfide cleavage, offers great potential for drug delivery.

Here we combine PCI with a covalently bound photosensitizer and the redox-labile disulfide-bridge approach to create highly potent multifunctional CMS nanoparticles acting as fluorimetric reporter vectors in different cell lines (Fig. 1).



**Fig. 1.** CMS-based endosome lysis detector system (A). CMS nanoparticles (NP), to which inactive protoporphyrin IX (PpIX) and the quenched dye ATTO 633 are covalently attached, are encapsulated in a lipid bilayer. Cells are incubated with NPs for at least 6 hours prior to a lysis experiment together with a fluid-phase marker, AlexaFluor Dextran 488. Endosome lysis is induced by activating PPIX with 405 nm light (B). The reactive oxygen species (ROS) produced by this reaction oxidize double bonds in the lipid tails in both the SLB surrounding the NPS, as well as of the endosome. Following membrane disruption, the reductive intracellular milieu cleaves the bonds between the quencher and ATTO 633, leading to localized fluorescence at the site of lysis.

Apart from the construction of a reporter system, it is useful to construct a model that can be used to understand variations in the lysis time distributions, especially the contributions of factors that could not be directly quantified, such as the PPIX load per NP. We have previously created a stochastic pi calculus model for describing the intracellular distribution of NPs and NP aggregates [52]. Here, we extended the model to include the PpIX-mediated endosome lysis reaction. The model is executed using SPiM

(<http://research.microsoft.com/en-us/projects/spim/>), developed by Andrew Philips and Luca Cardelli, at Microsoft Research, Cambridge[53, 54]. SPiM uses the Gillespie algorithm to describe the time evolution of the system by selecting the reaction that will occur, as well as its duration, with a probability proportional to the base rate of reaction and reactant availability [55]. Model results indicate that for a PpIX-dependent NP system, the lysis time distribution depends heavily on cell specific uptake parameters, rather than any inter-NP variability in PpIX load, provided that the PpIX load is not too low.

In this paper, we demonstrate that the use of this reporter system for generating time distributions of lysis events, coupled with an appropriate stochastic model, is a powerful method for evaluating and analyzing the efficiency of delivery vectors.

## Materials and Methods

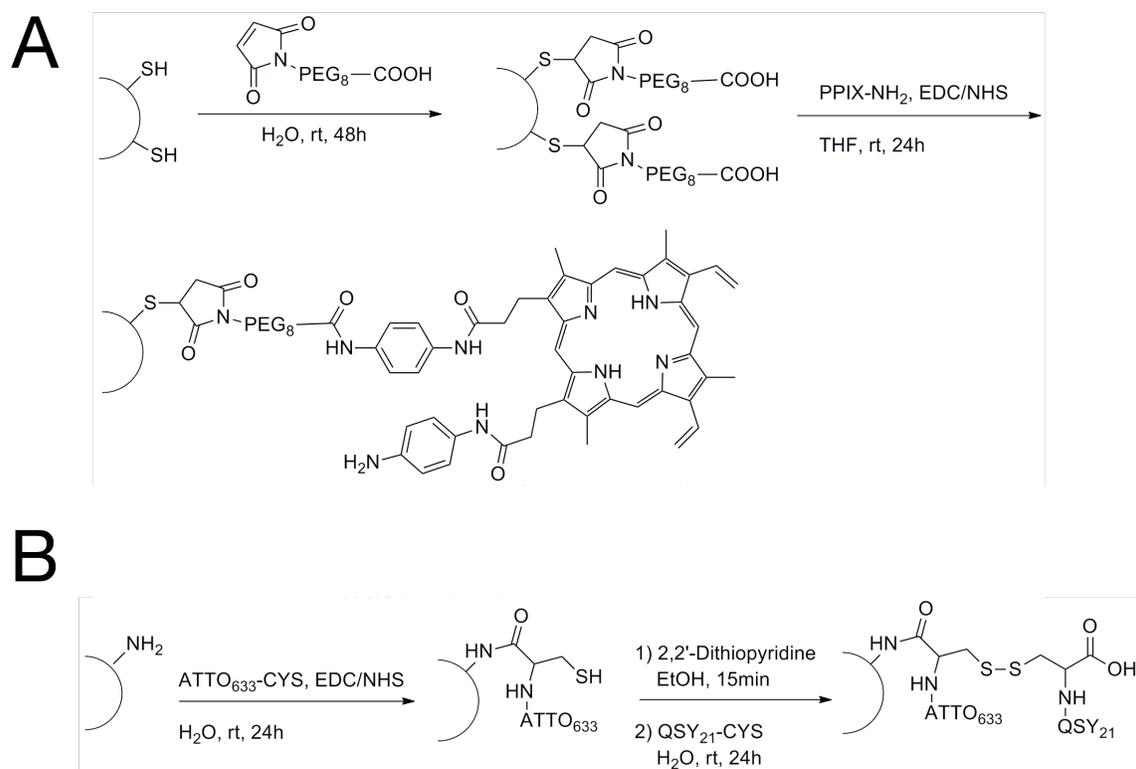
### *Multifunctional mesoporous core-shell silica nanoparticles*

The multifunctional core-shell CMS nanoparticles were synthesized following a previously reported procedure [6]. We use a so-called delayed co-condensation approach to selectively introduce different functionalities on the inner and outer particle surface. This synthesis employs tetraethylorthosilicate (TEOS) as silica source, 3-(aminopropyl)-triethoxysilane (APTES) and 3-(mercaptopropyl)-triethoxysilane as functionalized silica precursors, cetyltrimethylammonium chloride (CTAC) as structure directing agent. The polyalcohol triethanolamine (TEA) slows down the condensation rate of TEOS and leads to a small size distribution and radial growth of CMS. Functionalized CMS with 3-mercaptopropyl moieties in the shells of the particles and 3-aminopropyl groups in the core were synthesized according to this procedure giving the sample CMS\_NH<sub>2in</sub>-SH<sub>out</sub> after template removal (see Supporting Information for synthesis details and additional data). Post-synthesis modifications of the introduced functionalities are possible due to their reactive nature.

### *CMS based detector system*

Figure 2A shows the post-synthesis functionalization of the outer surface of the CMS particles. In the first step the shell was covered with a short, bifunctional polyethyleneglycol linker (maleimide-dPEG8-COOH, (1-maleimido-3-oxo-7,10,13,16,19,22,25,28-octaoxa-4-azahen-triacontan-31-oic acid), increasing monodispersity in water. The thiol moiety forms a stable bond with the maleimide functionalization of the PEG-linker in a Michael addition whereas the carboxy-functionality remains accessible for further modification. The amino-modified photosensitizer protoporphyrin-IX (PpIX-NH<sub>2</sub>) was synthesized according to a modified literature procedure [42]. This precursor was covalently attached to the carboxy-functions at the particle surface via an *N*-(3-dimethylaminopropyl)-*N*'-ethylcarbodiimide (EDC) assisted amidation to give the photosensitizer-modified sample CMS\_NH<sub>2in</sub>-PEG-PpIX<sub>out</sub>, as reported in detail in the Supporting Information.

We designed a redox-sensitive fluorimetric dye-quencher system for the CMS core (Figure 2B). In the first step the highly reactive ATTO 633-NHS ester was covalently attached to the amino group of cysteine (CysATTO 633). Subsequently, CysATTO 633 was reacted with the aminopropyl-functionality in an EDC assisted amidation to give the sample CMS\_CysATTO 633<sub>in</sub>-PEG-PPIX<sub>out</sub>. The thiol group of cysteine was activated with 2,2'-dithiopyridine, and cysteine-modified QSY21 (CysQSY21) was added. This resulted in a disulfide bridged dye-quencher system in the particle core and gave the final sample CMS\_CysATTO 633-CysQSY21<sub>in</sub>-PEG-PPIX<sub>out</sub>. Detailed information on the synthesis and labeling procedures are given in the Supporting Information.



**Fig 2. Shell (A) and core (B)-reactions of CMS-NPs.** Shell reactions: 1) PEGylation 2) PPIX-NH<sub>2</sub> attachment via EDC-amidation. Core reactions: 1) ATTO 633-CYS attachment via EDC-amidation 2) The 2,2'-dithiopyridine attachment to the SH-group of ATTO 633Cys is followed by the bonding of QSY21Cys via a redox-labile disulfide bridge.

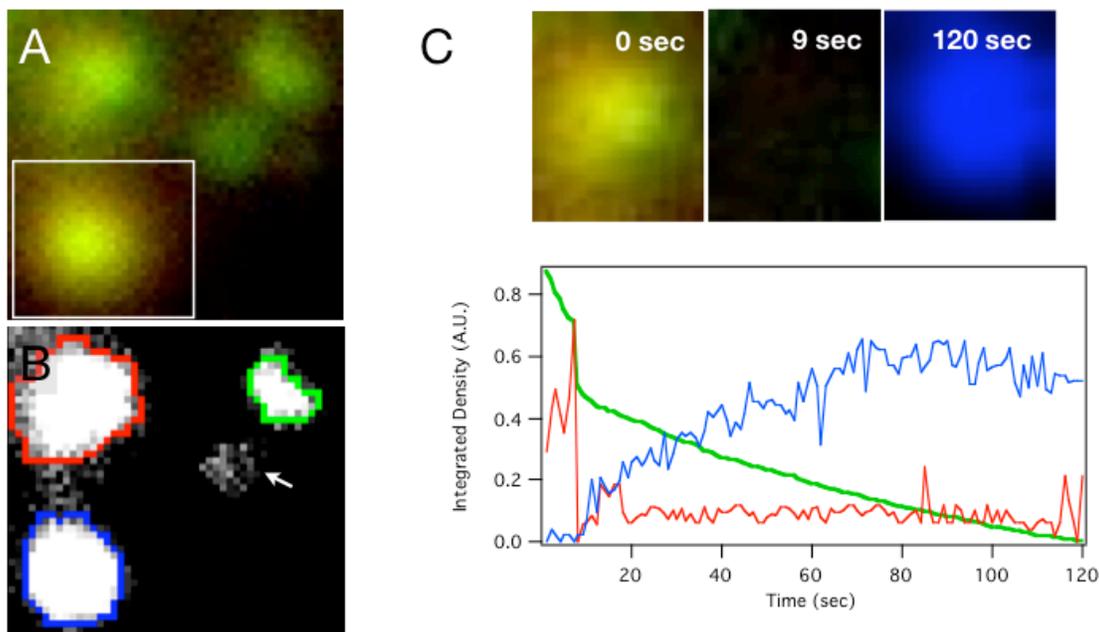
**Lipids.** 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC) and 1,2-Dioleoyl-sn-glycero-3-phospho-ethanolamine (DOPE), were chosen for the study. Both lipids bear a sufficiently high number of double bonds with which the singlet oxygen produced upon PPIX activation can react to destabilize the membrane structure [56]. Additionally, both lipids have been previously reported to form stable, fluid bilayers on the CMS core in its pure form [28, 57]. DOPE is known to possess fusogenic properties because of its ability to adopt an inverted hexagonal phase [58], influencing the curvature and structural integrity of a lipid bilayer [59], and can fuse with cell membranes better than DOPC [58-60]. CMS particles were covered with SLBs comprised of either pure DOPC (DOPC@CMS) or pure DOPE (DOPE@CMS) lipid membranes labelled with 1 mol % Texas Red 1,2-dihexadecanoyl-sn-glycero-3-phosphoethanolamine, triethylammonium salt (TR-DHPE).

**Supported lipid bilayer (SLB) deposition on CMS particles through solvent exchange.** SLBs were deposited on CMS particles as previously described [42]. Briefly, CMS NPs in solution are centrifuged at 13000 rpm for five minutes. The supernatant is discarded and the particles are redispersed in a 2.5 mg/ml lipid solution in 40% vol ethanol:60 % vol water; this solution was subsequently titrated with water until a 95% vol water content is reached. The solution is allowed to equilibrate for 15 minutes, in the process forming the SLB on individual particles [42]; the NPs are separated from excess lipid by centrifugation, and subsequently resuspended

in a smaller volume comprised of 50% vol complete cell culture medium:50 %vol water solution. In this alcohol-free environment, the SLB is fully stabilized. The aggregation of NPs in this solution is minimized by brief sonication followed by filtration.

*Uptake experiments.* DOPC@CMS particles were tested on non-phagocytic cell lines whose interactions with NPs were of interest: Beas2B, a human normal bronchial epithelium cell line; Huh7, a human hepatocarcinoma cell line; Renca-LacZ, a transformed murine renal carcinoma cell line; and 3T3, a mouse embryonic fibroblast cell line. Cells were plated to a density of  $0.75 \times 10^4$  cells/cm<sup>2</sup> on 8-well ibiTreat-coated Ibidi chambers 3 - 4 days before the start of the experiment. Cells were subsequently incubated with 50 µg/cm<sup>2</sup> DOPC@CMS over two-hour intervals up to eight hours, and washed with PBS five times post-incubation. External fluorescence was additionally quenched with a wash using 0.4 % w/w trypan blue [61, 62]. Cells were then fixed with 2 % formaldehyde in Leibovitz medium, reported to reduce cellular fluorescence [63], supplemented with 10% FCS for a total of two hours. Internalized NPs were visualized using fluorescence microscopy at 40x magnification, and recorded at an exposure time of 1 s. For each setup, eight random viewfields were chosen. Experiments were performed in triplicate. The relative number of particles taken up by each cell was approximated by taking the mean gray values across the area occupied by the cell. Analysis was done semi-automatically in ImageJ using an in-house program [64]. Control experiments were performed using 30 nm, fluorescein labeled Latex NPs (Sigma Aldrich).

*Lysis experiments.* For *in vitro* lysis experiments, cells were incubated for 12 hours with DOPC@CMS or DOPE@CMS functionalized with PpIX and ATTO 633-QSY21 to ensure completion of uptake. Lysis experiments were performed on 3T3 and RencaLacZ cells, which appear to exhibit the highest and lowest uptake characteristics for DOPC@CMS particles. AlexaFluor Dextran 488 (MW=10 kDa) was used as a fluid-phase marker for endocytosis. Following incubation, cells were washed with PBS five times post-incubation, or until most of the fluorescence from the AlexaFluor Dextran was removed; washing with 0.4 % w/w trypan blue was performed to quench any remaining external fluorescence. PpIX was activated by a one- or two –minute exposure to a 405 nm laser prior to imaging. Fluorescence images were taken every second with an exposure time of 200 ms over a period of two minutes per viewfield. NP and endosome fluorescence were monitored using the following excitation/emission filter sets: 488 nm/525 nm (AlexaFluor dextran); 560 nm/645 nm (Texas Red); and 640 nm/690 nm (ATTO 633) (Chroma Technology). A negative control composed of DOPC@CMS or DOPE@CMS functionalized with ATTO 633-QSY21, but not PPIX, was used in parallel with all setups. Single endosome lysis events were detected using the same ImageJ plugin described previously. Briefly, images are segmented based on the fluid-phase marker channel (**Fig. 3**). Discontinuities that occur in both the red and green channels were recorded as an endosome lysis event, and the time distribution of these events was collected for the different experimental setups. Apart from lysis times, the endosome area and NP colocalization intensity, based on the mean gray values in the red channel, are recorded.



**Fig. 3.** Single-endosome lysis time evaluation. Individual endosomes (A) are identified using automatic segmentation [64] performed on the green channel, which is able to identify three out of four endosomes; the arrow indicates an endosome that was not properly segmented, and that had to be manually added. Colors represent automatically-defined ROIs, and not RGB channels (B). ATTO 633 fluorescence is artificially colored blue. For each endosome, the lysis time is determined through the occurrence of discontinuities in the time course in the red and green channels, as well as an increase in fluorescence in the long red channel (C).

*Modelling.* To systematically evaluate the factors that influence lysis time distributions, we extended the generic stochastic pi calculus model that we have created for gold NP uptake and intracellular movement [65]. Specifically, we included PPIX and activated PPIX (Act\_PPIX) counts as attributes of each NP process; an activation reaction, coupled with preconditions for lysis in the form of a minimum amount of Act\_PPIX per endosome, also replaces the simple endosome lysis rate. We also redefine the aggregate size restrictions for 100 nm CMS NPs, instead of 4 nm gold NPs. Finally, we made uptake rates variable to represent cell specific responses. Each simulation was carried out with a starting value of 10000 NPs per cell; the PPIX load assigned to each NP was taken from a normal distribution with a mean of 50 PPIX per NP, and a width of 5. Sensitivity analysis was performed to identify the critical parameters of the model. We used code for SPiM v.0.05, which is run automatically using a Perl script that also permits parameter alteration. We omit the events downstream of endosomal escape, which we do not consider in this study.

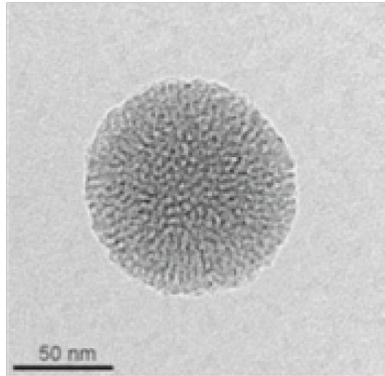
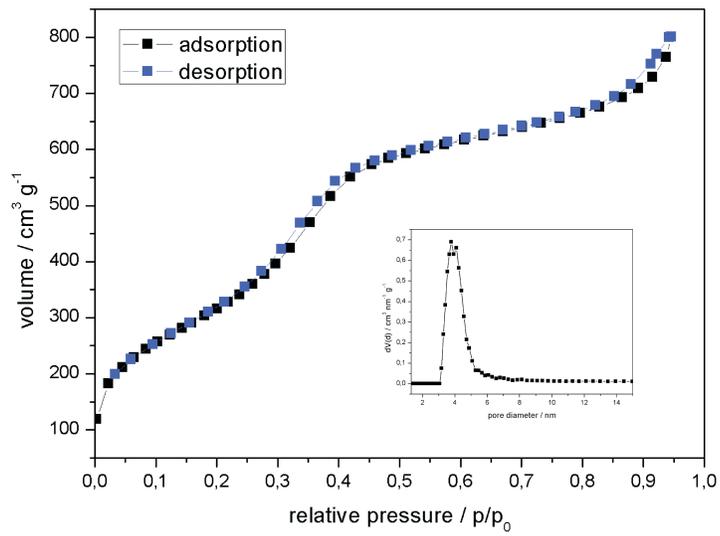
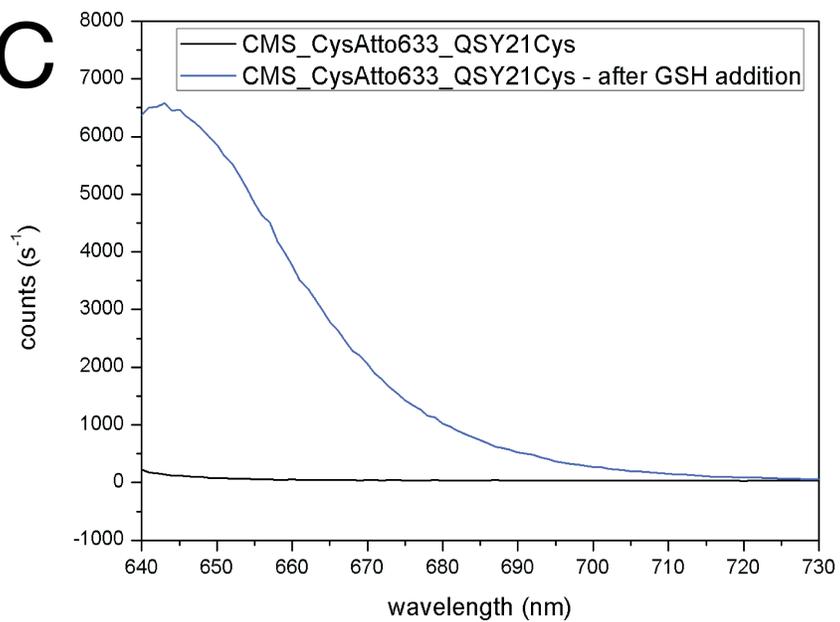
## Results and Discussion

### Reporter system characterization

The nitrogen sorption isotherm of the sample CMS-NH<sub>2</sub>core-SH<sub>shell</sub> is plotted in Fig. 4B upper panel. A Brunauer-Emmet-Teller (BET) surface area of 1160 m<sup>2</sup>g<sup>-1</sup> is calculated from the isotherm. The lower panel shows the pore size distribution derived from non-local density functional theory and shows a narrow pore size distribution (inset in Fig. 4B) of 3.7 nm. The

quality of the membrane was previously assessed using confocal microscopy and correlation spectroscopy, and was shown to prevent cargo release before it reaches the delivery target [28], or before controlled release is initiated. Here, we provide further characterization of the system using single-cell uptake experiments and single-endosome lysis experiments.

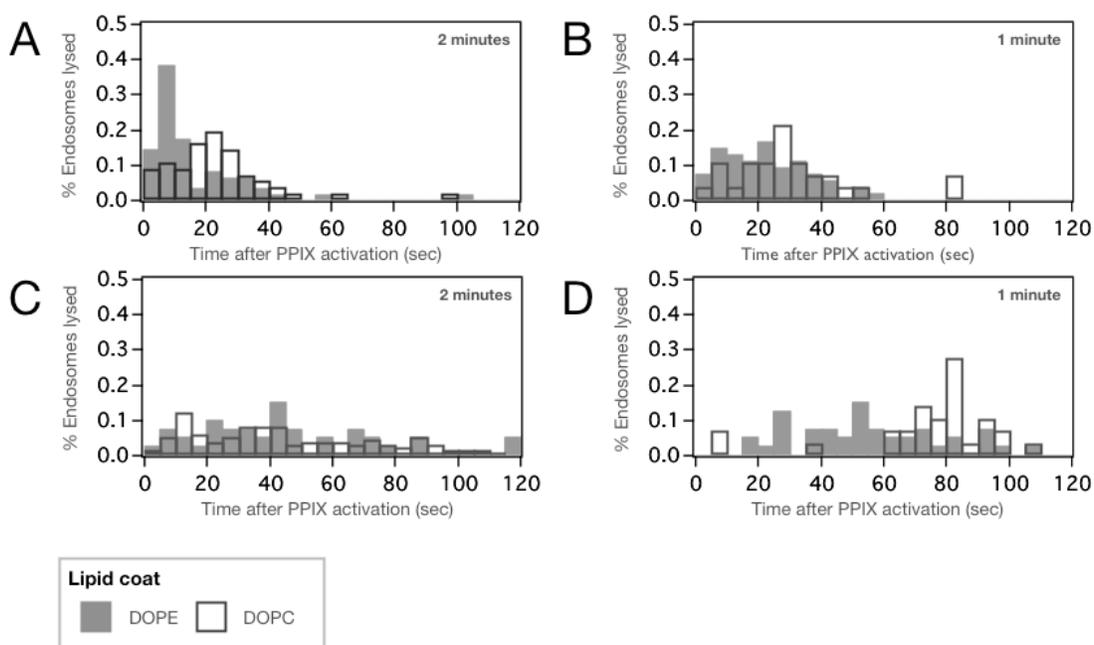
To demonstrate the efficiency of our disulfide bridged dye/quencher system in CMS\_CysATTO 633-QSY21Cys<sub>in</sub>-PEG-PpIX<sub>out</sub> the sample was investigated with fluorescence spectroscopy, Fig. 4C. The disulfide bridge between the two cysteine moieties is cleaved upon the addition of 10 mM L-glutathione which corresponds to the strength of the reductive milieu in the endosome. As a result, the covalently attached quencher CysQSY21 can diffuse out of the pores, resulting in a fluorescence of ATTO 633.

**A****B****C**

**Fig 4. CMS particle characterization.** Lipid-coated CMS NP imaged using Transmission electron microscopy (A). Nitrogen sorption isotherm (B, adsorption plotted in black, desorption in blue) and NLDFT pore size distribution (B, inset) of the sample CMS-NH<sub>2</sub>core-SH<sub>shell</sub>. Fluorescence emission spectra of the sample CMS\_CysATTO 633-QSY21CyS<sub>in</sub>-PEG-PpIX<sub>out</sub> (black) and 30 seconds after the addition of 10 mM GSH (blue) (C).

### Statistics from single-endosome PPIX-mediated lysis with different lipids, cell types, and PPIX activation times

We investigated the time distribution of endosome lysis as a function of PpIX activation time, lipid coat, and cell type. Our results generally show that DOPE@CMS exhibit marginally earlier lysis times than DOPC@CMS. An interesting observation is that the activation time appears to play a more prominent role in Renca-LacZ than in 3T3, as evinced by the bigger shift in the distribution when the activation time is decreased (**Fig. 5C-5D, Table 1**). Furthermore, the distribution of lysis times after a two-minute PPIX activation is considerably more spread than in 3T3, which could be an indication of cell-specific dependence of endosome lysis.



**Fig. 5.** Time distributions of endosome lysis events after PPIX-activation in 3T3 (A and B) and Renca-LacZ (C and D) cells. PPIX-functionalized DOPC@CMS and DOPE@CMS were incubated overnight, then activated for either one or two minutes. Results generally show marginally earlier lysis times for DOPE@CMS. It is interesting to note that lysis appears to be more concentrated towards earlier times for 3T3 cells than for RencaLacZ; furthermore, the shift in the lysis times when activation is shorter is less prominent in 3T3.

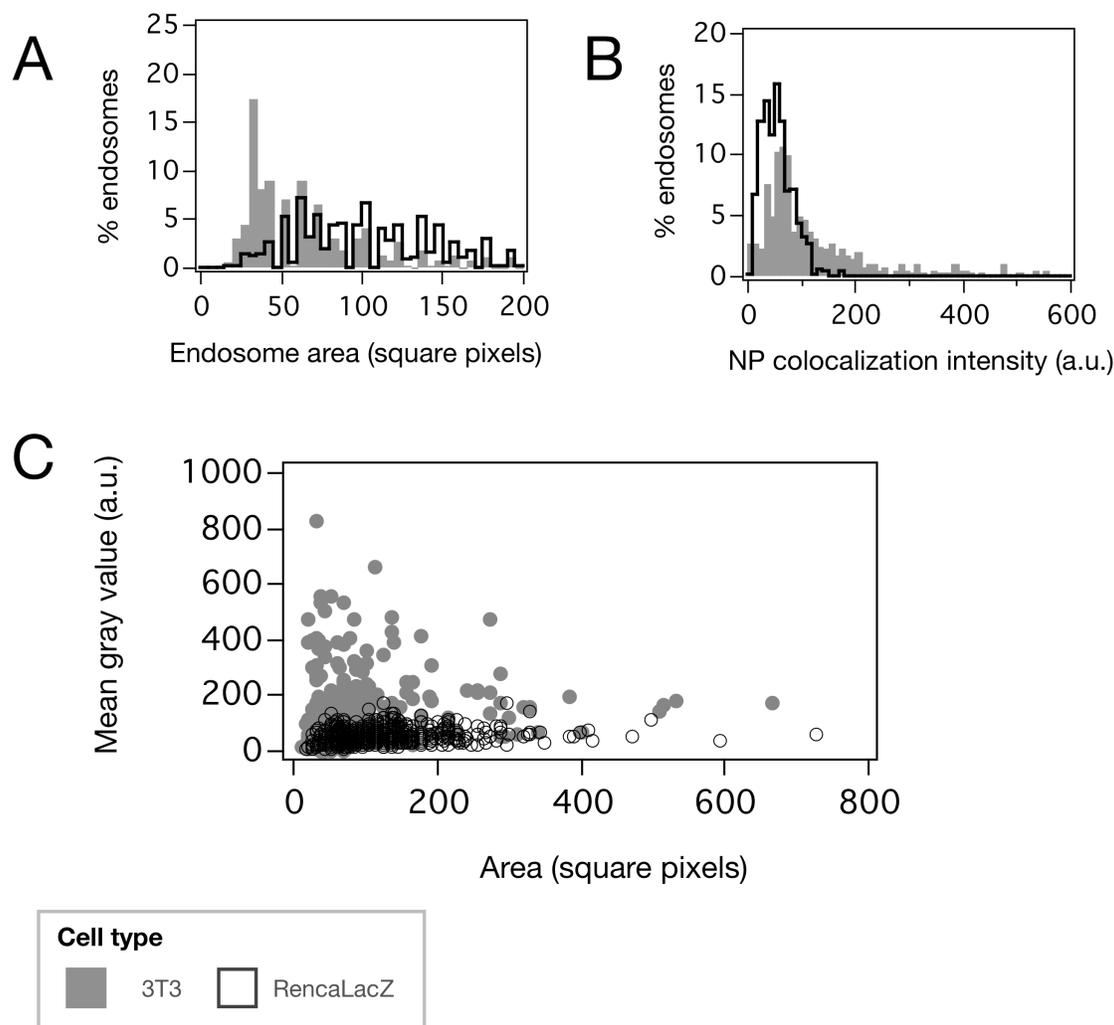
**Table 1.** Selected gauss fit parameters of endosomal escape data for DOPC@CMS and DOPE@CMS NPs in 3T3 and Renca-LacZ cultures

| Sample        | Activation Time (min) | Mean       | Width      |
|---------------|-----------------------|------------|------------|
| DOPC@CMS, 3T3 | 2                     | 17.4 ± 0.9 | 17.0 ± 1.5 |

|                      |   |                |                 |
|----------------------|---|----------------|-----------------|
| DOPC@CMS, 3T3        | 1 | 23.7 $\pm$ 1.8 | 17.2 $\pm$ 2.9  |
| DOPE@CMS, 3T3        | 2 | 5.3 $\pm$ 0.3  | 5.2 $\pm$ 0.4   |
| DOPE@CMS, 3T3        | 1 | 15.4 $\pm$ 1.6 | 24.0 $\pm$ 2.7  |
| DOPC@CMS, Renca-LacZ | 2 | 25.9 $\pm$ 7.3 | 39.3 $\pm$ 15.6 |
| DOPC@CMS, Renca-LacZ | 1 | 77.6 $\pm$ 1.9 | 13.6 $\pm$ 3.0  |
| DOPE@CMS, Renca-LacZ | 2 | 31.1 $\pm$ 5.3 | 28.0 $\pm$ 10.1 |
| DOPE@CMS, Renca-LacZ | 1 | 49.8 $\pm$ 5.0 | 35.0 $\pm$ 15.2 |

In order to check if there were cell-specific differences that may have influenced the lysis events, we analyzed the size distribution of endosomes, demarcated by the AlexaFluor Dextran 488. We also obtained the red channel intensity within the area occupied by the endosome as a rough approximation of the endosome load. Although the colocalization intensity distribution appears to have a similar range (**Fig. 6B**), the disparities in the endosome area are relatively marked (**Fig. 6A**), with Renca-LacZ cells displaying more size variation, and tending towards larger endosomes. When a correlation is made between individual endosomes and corresponding mean gray values (**Fig. 6C**), it becomes evident that 3T3 cells have a higher NP load per endosome on average.

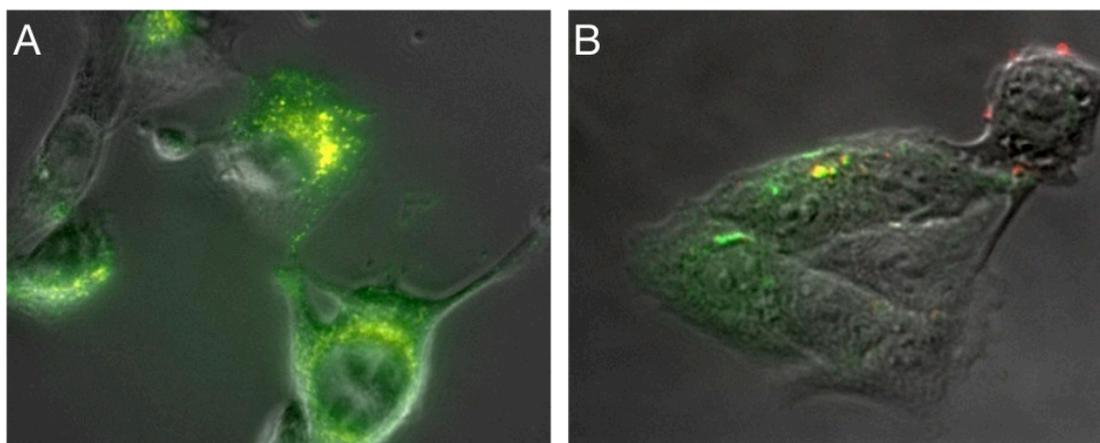
A possible reason for this is the higher expression and/or a longer lifetime of Rab5 in Renca-LacZ, a protein that controls endosome fusion, and consequently endosome size distribution [66]. Expression profiles of 3T3 cells and differentiated counterparts show an increase in Rab5 expression in differentiated cells (Guo et al. 2000); Rab5 has also been reported to be overexpressed in several carcinoma cell lines, including lung adenocarcinoma [67] and HeLa [68]. One way of clarifying the results would be to use a Rab5 tag to have an approximation of both its expression and lifetime in the cell types tested.



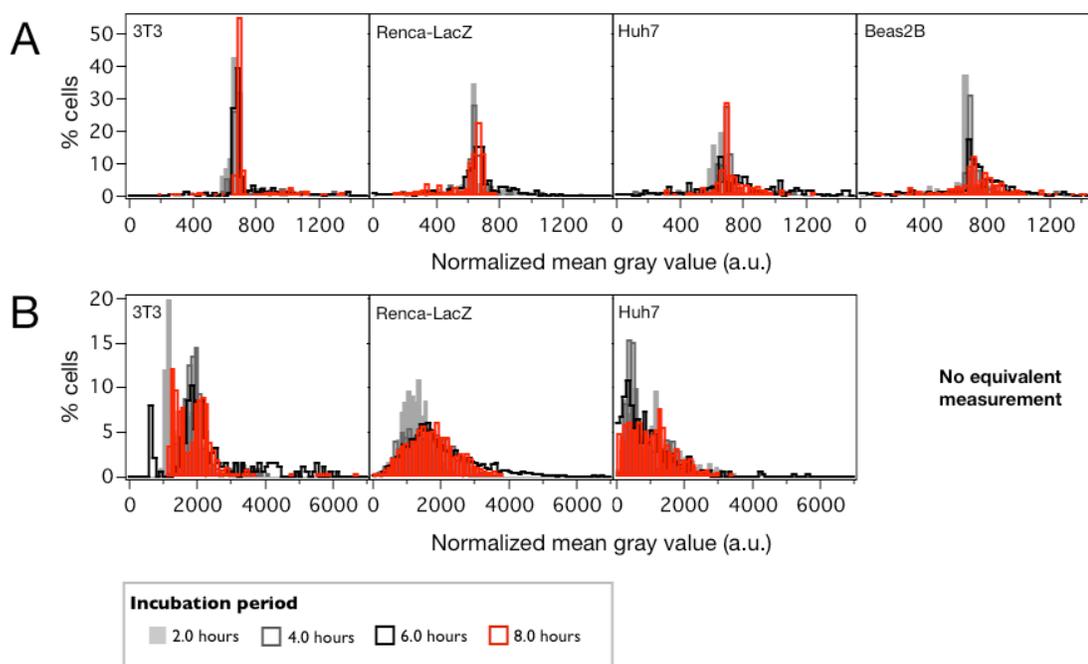
**Fig. 6.** Endosome size and colocalization intensity data for 3T3 and RencaLacZ cells. Endosomes are larger and exhibit higher size variation in RencaLacZ cells than in 3T3 cells (A), as well as a lower CMS reporter load (B, C). The variation in endosome size may be a result of differences in Rab5 expression.

### Cell-type dependence of NP uptake

Another interesting observation was that Renca-LacZ cells generally took up a smaller amount of CMS particles than 3T3 (**Fig. 7**). To confirm this, we conducted time-course uptake experiments using 3T3 and Renca cells, as well as two additional cell lines. From these experiments, we gathered information for between 500 and 1000 cells for each time point and each cell line. We ran a parallel experiment using 30 nm fluorescently-labeled Latex NPs as a form of control. Our results confirm that 3T3 cells take up more DOPC@CMS than Renca-LacZ on average (**Fig. 8, Table 2**). Additionally, the cell-to-cell uptake variation in Renca-LacZ cells appears to be at least twice of what could be expected from Renca-LacZ. **Table 2** summarizes selected parameters from Gaussian fits of the uptake data eight hours post-incubation. The cell-specific uptake features can also be observed with Latex NPs, with Renca-LacZ cells being able to take up the highest amount, followed by 3T3 and Huh7.



**Fig. 7.** Uptake in 3T3 cells (A) appears to be higher than in Renca-LacZ (B) after 12 hours of incubation with DOPC@CMS NPs. Note that individual cells of the same type appear to take up a widely variable numbers of NPs. NPs (red) were incubated together with Alexa-Fluor Dextran 488 (green). NPs inside endosomes appear as yellow spots. Images are taken at 40x.



**Fig. 8.** Cell type differences in reporter system uptake. Uptake of DOPC@CMS NPs (A) in 3T3, Renca-LacZ, Huh7 and Beas2B cells. The numbers of NPs taken up were approximated using the mean fluorescence intensity across a cell. A parallel experiment was performed using commercially available 30 nm latex nanoparticles (B) for 3T3, Renca-LacZ and Huh7 cells.

**Table 2.** Selected Gauss fit parameters of cell-type specific NP uptake profiles at eight hours post-incubation

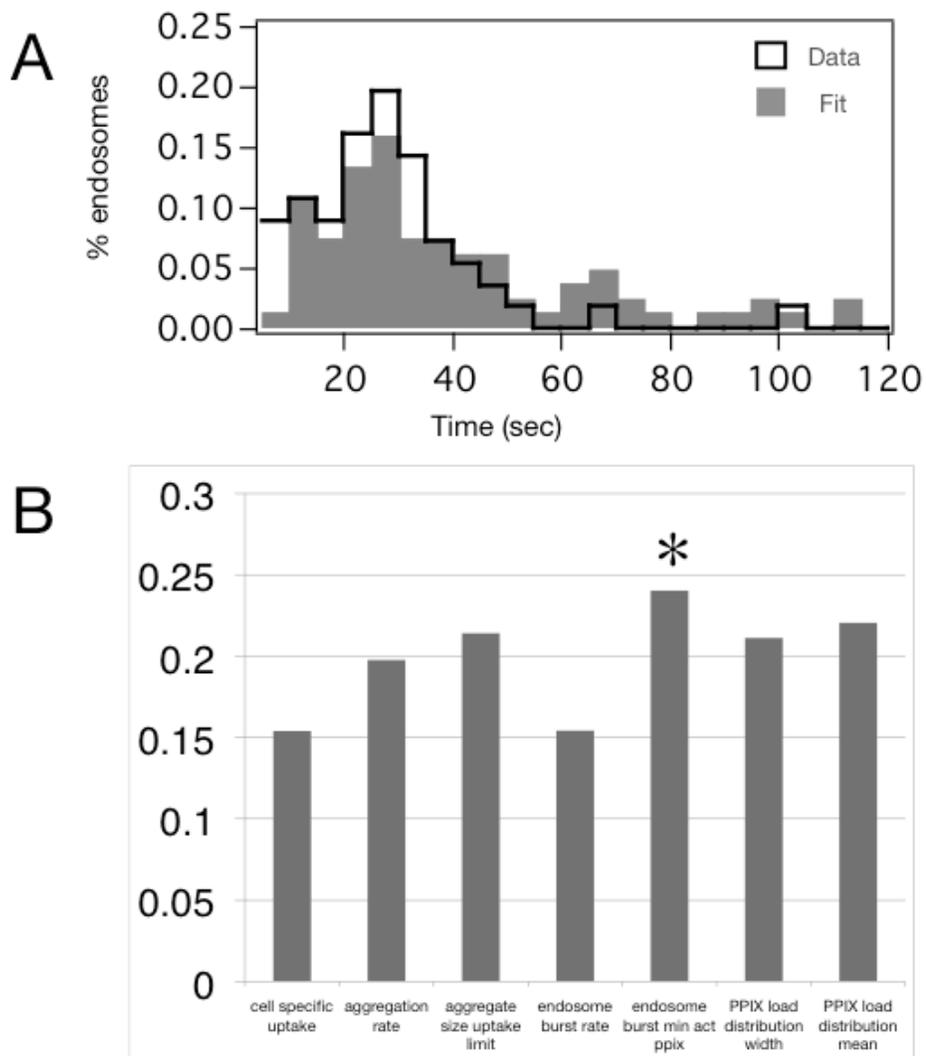
| Sample               | Mean        | Width      |
|----------------------|-------------|------------|
| DOPC@CMS, 3T3        | 675.7 ± 0.3 | 17.2 ± 0.2 |
| DOPC@CMS, Renca-LacZ | 644.5 ± 1.7 | 41.7 ± 2.4 |

|                   |                    |                    |
|-------------------|--------------------|--------------------|
| DOPC@CMS, Huh7    | $673.9 \pm 1.2$    | $20.9 \pm 1.1$     |
| DOPC@CMS, Beas2B  | $726.9 \pm 9.4$    | $132.5 \pm 14.3$   |
| Latex, 3T3        | $1314.1 \pm 29.9$  | $227.8 \pm 43.3$   |
| Latex, Renca-LacZ | $1603.54 \pm 19.3$ | $1065.7 \pm 31.6$  |
| Latex, Huh7       | $613.4 \pm 69.4$   | $1314.5 \pm 102.0$ |

It is not the first time that cell type uptake dependence for non-functionalized NPs was reported. Non-functionalized silver NPs, for instance, are taken up better by peripheral monocytes, but not by T-cell populations [69]. A recent paper also indicates that apart from uptake differences resulting from cell surface properties, there are also cell-type specific cytoplasmic and nuclear pore penetration constraints [70]. Our results, together with other recent reports on cell-type dependent uptake specificity, indicate the need to investigate this phenomenon more closely, and to take results as a consideration in NP design and dosage.

### Stochastic pi calculus model of endosomal escape

We extended the gold NP model to be able to systematically study the factors that influence endosomal escape. For a release-on-demand system such as PpIX, we expected a very narrow distribution of endosome lysis times. Considering the time window, which is two minutes post-PpIX activation, the spread is of course insignificant compared to non-release-on-demand systems. However, what was striking is the disparity of the distributions in 3T3 and Renca cells. In order to check how much of the spread is caused by noise inherent to the reporter system, we added PpIX and Act\_PpIX parameters to each NP, through which we can vary the PpIX load distribution per NP. As indicated in the methods section, we made the initial assumption that the NPs had a normally distributed PpIX load with a mean of 50, and a width of 5, which we changed in the course of the sensitivity analysis. For the first run, we also made the assumption that the binding and aggregation rates are the same as those used for gold NPs. Other parameters were also subjected to sensitivity analysis, results of which are summarized in **Fig. 9**. Interestingly, the model is most sensitive to the minimum number of PpIX required to burst an endosome. Although the current version of SPiM does not permit us to dynamically create endosomes that are assigned an unfixed size parameter (viz. taken from a distribution as in **Fig. 6**), this result is logical in that this minimum number is intimately linked with the NP load per endosome, as well as endosome size, which we have shown experimentally to be the main difference between the two cells.



**Fig. 9.** Fits of 3T3 data (A) and sensitivity analysis of the different model parameters (B) indicate that the minimum number of PPIX required to burst an endosome (indicated by an asterisk) is the most critical parameter in the model. Variations in NP batches could affect the results significantly as well, but our fit indicates an inter-NP variation of less than 10%. A more complete model that incorporates the effect of dynamically-assigned endosome size distribution, which is currently not possible in the existing simulator, would allow the results for Renca-LacZ to be fitted.

**Acknowledgements.** We would like to thank the Center for NanoScience (CeNS) for supporting this project. We also wish to acknowledge Dr. Valentina Cauda, Dr. Hanna Engelke, Dr. Anna Sauer, Dr. Axel Schlossbauer for significant discussions on handling and characterizing the CMS particles, and Dr. Thomas Ligon and Dr. Andrew Phillips for helpful discussions on SPiM. M.P.D. thanks the Deutscher Akademischer Austausch Dienst (DAAD) for her PhD grant.

## References

1. Douglas, K., *Toward Development of Artificial Viruses for Gene Therapy: A Comparative Evaluation of Viral and Non ...* Biotechnology Progress, 2008.
2. Varga, C.M., et al., *Quantitative comparison of polyethylenimine formulations and adenoviral vectors in terms of intracellular gene delivery processes*. Gene Therapy, 2005. **12**(13): p. 1023-1032.
3. Jin, H., D.A. Heller, and M.S. Strano, *Single-particle tracking of endocytosis and exocytosis of single-walled carbon nanotubes in NIH-3T3 cells*. Nano Lett., 2008. **8**(6): p. 1577-85.
4. Akita, H., et al., *Quantitative three-dimensional analysis of the intracellular trafficking of plasmid DNA transfected by a nonviral gene delivery system using confocal laser scanning microscopy*. Mol Ther, 2004. **9**(3): p. 443-51.
5. Chen, H.H., et al., *Quantitative Comparison of Intracellular Unpacking Kinetics of Polyplexes by a Model Constructed From Quantum Dot-FRET*. Molecular Therapy, 2008. **16**(2): p. 324-332.
6. Hama, S., et al., *Quantitative comparison of intracellular trafficking and nuclear transcription between adenoviral and lipoplex systems*. Mol Ther, 2006. **13**(4): p. 786-94.
7. Brabec, M., et al., *Opening of size-selective pores in endosomes during human rhinovirus serotype 2 in vivo uncoating monitored by single-organelle flow analysis*. Journal of Virology, 2005. **79**(2): p. 1008-16.
8. Brandenburg, B. and X. Zhuang, *Virus trafficking – learning from single-virus tracking*. Nat Rev Microbiol, 2007. **5**(3): p. 197-208.
9. Rust, M.J., et al., *Single-Virus Tracking in Live Cells*. Cold Spring Harbor Protocols, 2011. **2011**(9): p. pdb.top065623-pdb.top065623.
10. Chithrani, B.D. and W.C.W. Chan, *Elucidating the mechanism of cellular uptake and removal of protein-coated gold nanoparticles of different sizes and shapes*. Nano Lett., 2007. **7**(6): p. 1542-50.
11. Schwake, G., et al., *Predictive modeling of non-viral gene transfer*. Biotechnology and Bioengineering, 2009. **105**(4): p. 805-813.
12. Chia, S.U., Jun; Tamanoi, Fuyuhiko; Dunn, Bruce; Zink, Jeffrey, *Patterned Hexagonal Arrays of Living Cells in sol-Gel silica films*. J. Am. Chem. Soc., 2000. **122**: p. 6488 - 6489.
13. Radin, S., et al., *In vivo tissue response to resorbable silica xerogels as controlled-release materials*. Biomaterials, 2005. **26**(9): p. 1043-52.
14. Kortusuo, P., et al., *Silica xerogel as an implantable carrier for controlled drug delivery--evaluation of drug distribution and tissue effects after implantation*. Biomaterials, 2000. **21**(2): p. 193-8.
15. Möller, K., J. Kobler, and T. Bein, *Colloidal Suspensions of Nanometer-Sized Mesoporous Silica*. Adv. Funct. Mater., 2007. **17**(4): p. 605-612.
16. Kecht, J., A. Schlossbauer, and T. Bein, *Selective Functionalization of the Outer and Inner Surfaces in Mesoporous Silica Nanoparticles*. Chem. Mater., 2008. **20**(23): p. 7207-7214.
17. Cauda, V., et al., *Multiple core-shell functionalized colloidal mesoporous silica nanoparticles*. J. Am. Chem. Soc., 2009. **131**(32): p. 11361-70.
18. Muhammad, F., et al., *pH-Triggered controlled drug release from mesoporous silica nanoparticles via intracellular dissolution of ZnO nanolids*. J. Am. Chem. Soc., 2011. **133**(23): p. 8778-81.

19. Argyo, C., et al., *Heparin-Coated Colloidal Mesoporous Silica Nanoparticles Efficiently Bind to Antithrombin as an Anticoagulant Drug-Delivery System*. Chem. Eur. J., 2011. **18**(2): p. 428-432.
20. Rosenholm, J.M., et al., *Targeting of Porous Hybrid Silica Nanoparticles to Cancer Cells*. ACS Nano, 2009. **3**(1): p. 197-206.
21. Zhao, W., et al., *A glucose-responsive controlled release of insulin system based on enzyme multilayers-coated mesoporous silica particles*. Chem. Commun., 2011. **47**(33): p. 9459.
22. He, Q., et al., *The effect of PEGylation of mesoporous silica nanoparticles on nonspecific binding of serum proteins and cellular responses*. Biomaterials, 2009. **31**(6): p. 1085-1092.
23. Xia, T., et al., *Polyethyleneimine coating enhances the cellular uptake of mesoporous silica nanoparticles and allows safe delivery of siRNA and DNA constructs*. ACS Nano, 2009. **3**(10): p. 3273-86.
24. Kim, J., et al., *Multifunctional Uniform Nanoparticles Composed of a Magnetite Nanocrystal Core and a Mesoporous Silica Shell for Magnetic Resonance and Fluorescence Imaging and for Drug Delivery*. Angew. Chem. Int. Ed., 2008. **47**(44): p. 8438-8441.
25. Cauda, V., C. Argyo, and T. Bein, *Impact of different PEGylation patterns on the long-term bio-stability of colloidal mesoporous silica nanoparticles*. J. Mater. Chem., 2010. **20**(39): p. 8693.
26. Ogris, M., et al., *PEGylated DNA/transferrin-PEI complexes: reduced interaction with blood components, extended circulation in blood and potential for systemic gene delivery*. Gene Therapy, 1999. **6**(4): p. 595-605.
27. Cauda, V., A. Schlossbauer, and T. Bein, *Bio-degradation study of colloidal mesoporous silica nanoparticles: Effect of surface functionalization with organo-silanes and poly(ethylene glycol)*. Microporous and mesoporous materials, 2010. **132**(1-2): p. 60-71.
28. Cauda, V., et al., *Colchicine-Loaded Lipid Bilayer-Coated 50 nm Mesoporous Nanoparticles Efficiently Induce Microtubule Depolymerization upon Cell Uptake*. Nano Lett., 2010. **10**(7): p. 2484-2492.
29. Lee, C.-H., et al., *Intracellular pH-Responsive Mesoporous Silica Nanoparticles for the Controlled Release of Anticancer Chemotherapeutics*. Angew. Chem. Int. Ed., 2010. **49**(44): p. 8214-8219.
30. Lu, J., et al., *Mesoporous Silica Nanoparticles for Cancer Therapy: Energy-Dependent Cellular Uptake and Delivery of Paclitaxel to Cancer Cells*. Nanobiotechnology, 2007. **3**(2): p. 89-95.
31. Giri, S., B.G. Trewyn, and V.S.Y. Lin, *Mesoporous silica nanomaterial-based biotechnological and biomedical delivery systems*. Nanomedicine (Lond), 2007. **2**(1): p. 99-111.
32. Yang, P., S. Gai, and J. Lin, *Functionalized mesoporous silica materials for controlled drug delivery*. Chemical Society reviews, 2012. **41**(9): p. 3679.
33. Lai, C.-Y., et al., *A mesoporous silica nanosphere-based carrier system with chemically removable CdS nanoparticle caps for stimuli-responsive controlled release of neurotransmitters and drug molecules*. J. Am. Chem. Soc., 2003. **125**(15): p. 4451-9.
34. Meng, H., et al., *Autonomous in vitro anticancer drug release from mesoporous silica nanoparticles by pH-sensitive nanovalves*. J. Am. Chem. Soc., 2010. **132**(36): p. 12690-7.

35. Khashab, N.M., et al., *Redox- and pH-Controlled Mechanized Nanoparticles*. Eur. J. Org. Chem., 2009. **2009**(11): p. 1669-1673.
36. Ambrogio, M.W., et al., *Snap-top nanocarriers*. Org Lett, 2010. **12**(15): p. 3304-7.
37. Radu, D.R., et al., *A polyamidoamine dendrimer-capped mesoporous silica nanosphere-based gene transfection reagent*. J. Am. Chem. Soc., 2004. **126**(41): p. 13216-7.
38. Gan, Q., et al., *A magnetic, reversible pH-responsive nanogated ensemble based on Fe<sub>3</sub>O<sub>4</sub> nanoparticles-capped mesoporous silica*. Biomaterials, 2011. **32**(7): p. 1932-42.
39. Mal, N.F., M; Tanaka, Y., *Photocontrolled reversible release of guest molecules from coumarin-modified mesoporous silica*. Nature, 2003. **421**(6921): p. 350-353.
40. Schlossbauer, A., J. Kecht, and T. Bein, *Biotin-Avidin as a Protease-Responsive Cap System for Controlled Guest Release from Colloidal Mesoporous Silica*. Angew. Chem. Int. Ed., 2009. **48**(17): p. 3092-3095.
41. Schlossbauer, A., et al., *A Programmable DNA-Based Molecular Valve for Colloidal Mesoporous Silica*. Angew. Chem. Int. Ed., 2010. **49**(28): p. 4734-4737.
42. Hohner, A.O., M.P.C. David, and J.O. Rädler, *Controlled solvent-exchange deposition of phospholipid membranes onto solid surfaces*. Biointerphases, 2010. **5**(1): p. 1-8.
43. Liu, J., et al., *Porous nanoparticle supported lipid bilayers (protocells) as delivery vehicles*. J. Am. Chem. Soc., 2009. **131**(4): p. 1354-5.
44. Duncanson, W.J., et al., *Targeted binding of PLA microparticles with lipid-PEG-tethered ligands*. Biomaterials, 2007. **28**(33): p. 4991-9.
45. Liu, J., et al., *Electrostatically mediated liposome fusion and lipid exchange with a nanoparticle-supported bilayer for control of surface charge, drug containment, and delivery*. J. Am. Chem. Soc., 2009. **131**(22): p. 7567-9.
46. Berg, K., et al., *Photochemical Internalization (PCI): A Technology for Drug Delivery*. Methods in Molecular Biology, 2010. **635**: p. 133 - 145.
47. Bruin, K.G.d., et al., *Dynamics of photoinduced endosomal release of polyplexes*. Journal of Controlled Release, 2008. **130**(2): p. 175-182.
48. Yang, Y., et al., *Lipid coated mesoporous silica nanoparticles as photosensitive drug carriers*. Phys. Chem. Chem. Phys., 2010. **12**(17): p. 4418.
49. Ma, Y., et al., *pH-responsive mitoxantrone (MX) delivery using mesoporous silica nanoparticles (MSN)*. J. Mater. Chem., 2011. **21**(26): p. 9483.
50. Schlossbauer, A., et al., *pH-Responsive Release of Acetal-Linked Melittin from SBA-15 Mesoporous Silica*. Angew. Chem. Int. Ed., 2011. **50**(30): p. 6828-6830.
51. Sauer, A.M., et al., *Role of Endosomal Escape for Disulfide-Based Drug Delivery from Colloidal Mesoporous Silica Evaluated by Live-Cell Imaging*. Nano Lett., 2010. **10**(9): p. 3684-3691.
52. Dobay, M.P., et al., *Modeling nanoparticle uptake and intracellular distribution using stochastic process algebras*. J Nanopart Res, 2012. **14**: p. 821 - 833.
53. Cardelli, L. and A. Phillips, *A correct abstract machine for the stochastic pi-calculus*. Procs. BioConcur, 2004. **90**.
54. Phillips, A. and L. Cardelli, *A Correct Abstract Machine for the Stochastic Pi-calculus*. Concurrent Models in Molecular Biology, August 2004.
55. Gillespie..., D., *Exact stochastic simulation of coupled chemical reactions*. The Journal of Physical Chemistry, 1977.
56. Maziere, J.C., et al., *Cellular uptake and photosensitizing properties of anticancer porphyrins in cell membranes and low and high density lipoproteins*. J Photochem Photobiol B, Biol, 1990. **6**(1-2): p. 61-8.

57. Khalil, I.A., et al., *Uptake pathways and subsequent intracellular trafficking in nonviral gene delivery*. *Pharmacol Rev*, 2006. **58**(1): p. 32-45.
58. Hafez, I.M. and P.R. Cullis, *Roles of lipid polymorphism in intracellular delivery*. *Advanced Drug Delivery Reviews*, 2001. **47**: p. 139 -- 148.
59. Koltover, I., et al., *An Inverted Hexagonal Phase of Cationic Liposome-DNA Complexes Related to DNA Release and Delivery*. *Science*, 1998. **281**(5373): p. 78-81.
60. Bailey, A.L. and P.R. Cullis, *Membrane fusion with cationic liposomes: effects of target membrane lipid composition*. *Biochemistry*, 1997. **36**(7): p. 1628-1634.
61. Rejman, J., V. Oberle, and I. Zuhorn..., *Size-dependent internalization of particles via the pathways of clathrin-and caveolae-mediated endocytosis*. *Biochemical Journal*, 2004.
62. Hed, J., et al., *The use of fluorescence quenching in flow cytometry to measure the attachment and ingestion phases in phagocytosis in peripheral blood without prior cell separation*. *J Immunol Methods*, 1987. **101**(1): p. 119-125.
63. Xiao, J., *Handbook of Single-Molecule Biophysics*, ed. P. Hinterdorfer and A.v. Oijen 2009, New York: Springer.
64. Youssef, S., S. Gude, and J.O. Radler, *Automated tracking in live-cell time-lapse movies*. *Integrative Biology*, 2011. **3**(11): p. 1095-1101.
65. Dobay, M.P., et al., *How many trimers? Modeling influenza virus fusion yields a minimum aggregate size of six trimers, three of which are fusogenic*. *Mol. BioSyst.*, 2011. **7**(10): p. 2741-2749.
66. Li, G. and H. Qian, *Kinetic timing: a novel mechanism that improves the accuracy of GTPase timers in endosome fusion and other biological processes*. *Traffic*, 2002. **3**(4): p. 249-55.
67. Li, Y., et al., *Over-expression of the RAB5 gene in human lung adenocarcinoma cells with high metastatic potential*. *Chin Med Sci J*, 1999. **14**(2): p. 96-101.
68. Duclos, S., et al., *Rab5 regulates the kiss and run fusion between phagosomes and endosomes and the acquisition of phagosome leishmanicidal properties in RAW 264.7 macrophages*. *J Cell Science*, 2000. **113**: p. 3531 - 3541.
69. Greulich, C., et al., *Cell type-specific responses of peripheral blood mononuclear cells to silver nanoparticles*. *Acta Biomaterialia*, 2011. **7**(9): p. 3505-3514.
70. Williams, Y., et al., *Probing Cell-Type-Specific Intracellular Nanoscale Barriers Using Size-Tuned Quantum Dots*. *Small*, 2009. **5**(22): p. 2581-2588.



# Bibliography

- [1] Pier Luigi Luisi, Tereza Pereira de Souza, and Pasquale Stano. Vesicle behavior: in search of explanations. *The Journal of Physical Chemistry B*, 112(46):14655–64, Nov 2008.
- [2] Devika Soundara Manickam and David Oupický. Multiblock reducible copolypeptides containing histidine-rich and nuclear localization sequences for gene delivery. *Bioconjugate Chem.*, 17(6):1395–403, Jan 2006.
- [3] Robby A Petros and Joseph M Desimone. Strategies in the design of nanoparticles for therapeutic applications. *Nature reviews Drug discovery*, 9(8):615, Jul 2010.
- [4] Tomasz Puzyn, Danuta Leszczynska, and Jerzy Leszczynski. Toward the development of "nano-qsars": advances and challenges. *Small*, 5(22):2494–509, Nov 2009.
- [5] Tomasz Puzyn, Bakhtiyor Rasulev, Agnieszka Gajewicz, Xiaoke Hu, Thabitha P Dasari, Andrea Michalkova, Huey-Min Hwang, Andrey Toropov, Danuta Leszczynska, and Jerzy Leszczynski. Using nano-qsar to predict the cytotoxicity of metal oxide nanoparticles. *Nature Nanotechnology*, pages 1–4, Feb 2011.
- [6] Denis Fourches, Dongqiuye Pu, Carlos Tassa, Ralph Weissleder, Stanley Y Shaw, Russell J Mumper, and Alexander Tropsha. Quantitative nanostructure-activity relationship modeling. *ACS Nano*, 4(10):5703–12, Oct 2010.
- [7] Christie Sayes and Ivan Ivanov. Comparative study of predictive computational models for nanoparticle-induced cytotoxicity. *Risk Analysis*, 30(11):1723–1734, Jun 2010.
- [8] Jianzhong Liu and Anton J Hopfinger. Identification of possible sources of nanotoxicity from carbon nanotubes inserted into membrane bilayers using membrane interaction quantitative structure–activity relationship analysis. *Chem Res Toxicol*, 21(2):459–66, Feb 2008.
- [9] Huw D Summers, Paul Rees, Mark D Holton, M Rowan Brown, Sally C Chappell, Paul J Smith, and Rachel J Errington. Statistical analysis of nanoparticle dosing in a dynamic cellular system. *Nature Nanotechnology*, 6:1–5, Feb 2011.

- [10] Kate Dunster, Ban Hock Toh, and John W Sentry. Early endosomes, late endosomes, and lysosomes display distinct partitioning strategies of inheritance with similarities to golgi-derived membranes. *Eur J Cell Biol*, 81(3):117–24, Mar 2002.
- [11] María A Rujano, Floris Bosveld, Florian A Salomons, Freark Dijk, Maria A.W.H Van Waarde, Johannes J.L Van Der Want, Rob A.I De Vos, Ewout R Brunt, Ody C.M Sibon, and Harm H Kampinga. Polarised asymmetric inheritance of accumulated protein damage in higher eukaryotes. *Plos Biol*, 4(12):e417, Jan 2006.
- [12] Aakanksha Singhvi and Gian Garriga. Asymmetric divisions, aggresomes and apoptosis. *Trends in Cell Biology*, 19(1):1–7, Jan 2009.
- [13] D A Nauman and C R Bertozzi. Kinetic parameters for small-molecule drug delivery by covalent cell surface targeting. *Biochimica et biophysica acta*, 1568(2):147–54, Dec 2001.
- [14] Huajian Gao, Wendong Shi, and Lambert B Freund. Mechanics of receptor-mediated endocytosis. *Proc Natl Acad Sci USA*, 102(27):9469–74, Jul 2005.
- [15] Hongyan Yuan, Ju Li, Gang Bao, and Sulin Zhang. Variable nanoparticle-cell adhesion strength regulates cellular uptake. *Phys Rev Lett*, 105(13):138101, Sep 2010.
- [16] Hongyan Yuan and Zhang Sulin. Effects of particle size and ligand density on the kinetics of receptor-mediated endocytosis of nanoparticles. *Appl. Phys. Lett.*, 96:033704–1 – 033704–3, 2010.
- [17] Y Sidorenko and U Reichl. Structured model of influenza virus replication in mdck cells. *Biotechnology and Bioengineering*, Jan 2004.
- [18] G Schwake, S Youssef, JT Kuhr, S Gude, MP David, E Mendoza, E Frey, and JO Rädler. Predictive modeling of non-viral gene transfer. *Biotechnology and Bioengineering*, 105(4):805–813, 2009.
- [19] C M Varga, N C Tedford, M Thomas, A M Klivanov, 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, Jul 2005.
- [20] Bree B Aldridge, John M Burke, Douglas A Lauffenburger, and Peter K Sorger. Physicochemical modelling of cell signalling pathways. *Nat Cell Biol*, 8(11):1195–203, Nov 2006.
- [21] Thibault Lagache, Olivier Danos, and David Holcman. Modeling the step of endosomal escape during cell infection by a nonenveloped virus. *Biophysj*, 102(5):980–989, Mar 2012.

- [22] Raul E Cachau, Fernando D Gonzalez-Nilo, Oscar N Ventura, and Martin J Fritts. In-silico nanobio-design. a new frontier in computational biology. *Current topics in medicinal chemistry*, 7(15):1537–40, Jan 2007.
- [23] Mingguang Li, Khuloud T Al-Jamal, Kostas Kostarelos, and Joshua Reineke. Physiologically based pharmacokinetic modeling of nanoparticles. *ACS Nano*, 4(11):6303–17, Nov 2010.
- [24] A Regev and E Shapiro. The pi calculus as an abstraction for biomolecular systems. *Nat. Comp. Ser.*, pages 219–266, Jan 2004.
- [25] Carlee E Ashley, Eric C Carnes, Genevieve K Phillips, David Padilla, Paul N Durfee, Page A Brown, Tracey N Hanna, Juwen Liu, Brandy Phillips, Mark B Carter, Nick J Carroll, Xingmao Jiang, Darren R Dunphy, Cheryl L Willman, Dimiter N Petsev, Deborah G Evans, Atul N Parikh, Bryce Chackerian, Walker Wharton, David S Peabody, and C. Jeffrey Brinker. The targeted delivery of multicomponent cargos to cancer cells by nanoporous particle-supported lipid bilayers. *Nature Materials*, 10(5):389–397, May 2011.
- [26] E Wagner, C Plank, K Zatloukal, M Cotten, and ML Birnstiel. Influenza virus hemagglutinin ha-2 n-terminal fusogenic peptides augment gene transfer by transferrin-polylysine-dna complexes: toward a synthetic virus-like gene-transfer vehicle. *Proc Natl Acad Sci USA*, 89(17):7934, 1992.
- [27] K Douglas. Toward development of artificial viruses for gene therapy: A comparative evaluation of viral and non . . . . *Biotechnology Progress*, Jan 2008.
- [28] Catarina Madeira, Luís M S Loura, Maria R Aires-Barros, and Manuel Prieto. Fluorescence methods for lipoplex characterization. *BBA - Biomembranes*, 1808(11):2694–2705, Aug 2011.
- [29] V F Van Tendeloo, C Van Broeckhoven, and Z N Berneman. Gene therapy: principles and applications to hematopoietic cells. *Leukemia*, 15(4):523–44, Apr 2001.
- [30] Michael L Edelstein, Mohammad R Abedi, and Jo Wixon. Gene therapy clinical trials worldwide to 2007—an update. *The Journal of Gene Medicine*, 9(10):833–842, Oct 2007.
- [31] Michael L Edelstein, Mohammad R Abedi, Jo Wixon, and Richard M Edelstein. Gene therapy clinical trials worldwide 1989–2004—an overview. *The Journal of Gene Medicine*, 6(6):597–602, May 2004.
- [32] Soma Patnaik, Mohammed Arif, Atul Pathak, Naresh Singh, and K C Gupta. Pei-alginate nanocomposites: efficient non-viral vectors for nucleic acids. *International Journal of Pharmaceutics*, 385(1-2):194–202, Jan 2010.

- [33] Ethlenn V. B Gaal, Ronald S Oosting, Roel Eijk, Marta Bakowska, Dries Feyen, Robbert Jan Kok, Wim E Hennink, Daan J. A Crommelin, and Enrico Mastrobatista. Dna nuclear targeting sequences for non-viral gene delivery. *Pharmaceutical Research*, 28(7):1707–1722, Jul 2011.
- [34] Gabriele D Schmidt-Wolf and Ingo G.H Schmidt-Wolf. Non-viral and hybrid vectors in human gene therapy: an update. *Trends in molecular medicine*, 9(2):67–72, Feb 2003.
- [35] A Mountain. Gene therapy: the first decade. *Trends in Biotechnology*, 18(3):119–28, Mar 2000.
- [36] Zinnia P Parra-Guillén, Gloria González-Aseguinolaza, Pedro Berraondo, and Iñaki F Trocóniz. Gene therapy: A pharmacokinetic/pharmacodynamic modelling overview. *Pharmaceutical Research*, 27(8):1487–1497, Aug 2010.
- [37] E M Campbell and T J Hope. Gene therapy progress and prospects: Viral trafficking during infection. *Gene Therapy*, 12(18):1353–1359, Sep 2005.
- [38] Ming Je Hsieh, Paul J White, and Colin W Pouton. Interaction of viruses with host cell molecular motors. *Curr Opin Biotechnol*, 21(5):633–639, Oct 2010.
- [39] C Pichon, L Billiet, and P Midoux. Chemical vectors for gene delivery: uptake and intracellular trafficking. *Curr Opin Biotechnol*, 2010.
- [40] L K Medina-Kauwe, J Xie, and S Hamm-Alvarez. Intracellular trafficking of nonviral vectors. *Gene Therapy*, 12(24):1734–1751, Dec 2005.
- [41] K Comfort, E Maurer, and L Braydich-Stolle. . . . Interference of silver, gold, and iron oxide nanoparticles on egf signal transduction in epithelial cells. *ACS Nano*, Jan 2011.
- [42] Jasmin Fisher and Thomas A Henzinger. Executable cell biology. *Nat Biotechnol*, 25(11):1239, Nov 2007.
- [43] Robin Milner. *Communicating and Mobile Systems: the Pi-Calculus*. Cambridge University Press, 1st edition, June 1999.
- [44] M David, J Bantang, and E Mendoza. A projective brane calculus with activate, bud and mate as primitive actions. *Transactions on Computational Systems Biology XI*, pages 164–186, 2009.
- [45] L Cardelli. Brane calculi. *CMSB*, 4:257–278, 2004.
- [46] A Regev, W Silverman, and E Shapiro. Representation and simulation of biochemical processes using the pi-calculus process algebra. *Pac Symp Biocomput*, pages 459–70, Jan 2001.

- [47] C Priami. Stochastic pi-calculus. *The Computer Journal*, 38:578–589, 1995.
- [48] Dennis Yq Wang, Luca Cardelli, Andrew Phillips, Nir Piterman, and Jasmin Fisher. Computational modeling of the egfr network elucidates control mechanisms regulating signal dynamics. *BMC Syst Biol*, 3(1):118, Jan 2009.
- [49] N Owens, J Timmis, A Greensted, and A Tyrrell. Modelling the tunability of early t cell signalling events. *Lecture Notes in Computer Science*, 5132:12–23, 2008.
- [50] L Cardelli and A Phillips. A correct abstract machine for the stochastic pi-calculus. *Procs. BioConcur*, 90, 2004.
- [51] A Bittig and A Uhrmacher. Spatial modeling in cell biology at multiple levels. *Simulation Conference (WSC), Proceedings of the 2010 Winter*, pages 608 – 619, 2010.
- [52] V Danos, J Feret, J Fontana, R Harmer, and J Krivine. Rule-based modelling of cellular signalling. *Proceedings of the 18 th International Conference on Concurrency Theory (CONCUR'07), Lecture Notes in Computer Science*, 4703:17 – 41, 2007.
- [53] William S Hlavacek, James R Faeder, Michael L Blinov, Richard G Posner, Michael Hucka, and Walter Fontana. Rules for modeling signal-transduction systems. *Sci STKE*, 2006(344):re6, Jul 2006.
- [54] Kanae Oda, Yukiko Matsuoka, Akira Funahashi, and Hiroaki Kitano. A comprehensive pathway map of epidermal growth factor receptor signaling. *Molecular Systems Biology*, 1(1):E1–E17, May 2005.
- [55] D Gillespie. . . . Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, Jan 1977.
- [56] M Gibson and Jehoshua Bruck. Efficient exact stochastic simulation of chemical systems with many species and many channels. *J Phys Chem A*, 104:1876 – 1889, 2000.
- [57] Y Cao, D Gillespie, and L Petzold. The slow-scale stochastic simulation algorithm. *The Journal of Chemical Physics*, Jan 2005.
- [58] Y Cao and L Petzold. Slow-scale tau-leaping method. *Computer Methods in Applied Mechanics and Engineering*, Jan 2008.
- [59] Y Cao, D Gillespie, and L Petzold. Avoiding negative populations in explicit poisson tau-leaping. *The Journal of Chemical Physics*, Jan 2005.
- [60] V Danos, S Pradalier, and ENS Cachan. Projective brane calculus. *Computational Methods in Systems Biology: International Conference CMSB 2004, Paris, France, May 26-28, 2004: Revised Selected Papers*, page 134, 2005.

- [61] A Copic, C. F Latham, M. A Horlbeck, J. G D'arcangelo, and E. A Miller. Er cargo properties specify a requirement for copii coat rigidity mediated by sec13p. *Science*, 335(6074):1359–1362, Mar 2012.
- [62] H Bouwmeester, J Poortman, R.J Peters, E Wijma, E Kramer, S Makama, K Puspitaninganindita, H.J.P Marvin, A.A.C.M Peijnenburg, and P.J.M Hendriksen. Characterization of translocation of silver nanoparticles and effects on whole-genome gene expression using an in vitro intestinal epithelium co-culture model. *ACS Nano*, 2011.
- [63] M Hucka, A Finney, H. M Sauro, H Bolouri, J. C Doyle, H Kitano, , The Rest Of The Sbml Forum, A. P Arkin, B. J Bornstein, D Bray, A Cornish-Bowden, A. A Cuellar, S Dronov, E. D Gilles, M Ginkel, V Gor, I. I Goryanin, W. J Hedley, T. C Hodgman, J.-H Hofmeyr, P. J Hunter, N. S Juty, J. L Kasberger, A Kremling, U Kummer, N Le Novere, L. M Loew, D Lucio, P Mendes, E Minch, E. D Mjolsness, Y Nakayama, M. R Nelson, P. F Nielsen, T Sakurada, J. C Schaff, B. E Shapiro, T. S Shimizu, H. D Spence, J Stelling, K Takahashi, M Tomita, J Wagner, and J Wang. The systems biology markup language (sbml): a medium for representation and exchange of biochemical network models. *Bioinformatics*, 19(4):524–531, Mar 2003.
- [64] C Eccher and C Priami. Design and implementation of a tool for translating sbml into the biochemical stochastic -calculus. *Bioinformatics*, 22(24):3075–3081, Dec 2006.
- [65] Enrico Burello and Andrew Worth. Computational nanotoxicology: Predicting toxicity of nanoparticles. *Nature Publishing Group*, 6(3):138–139, Mar 2011.
- [66] Ralph A Sperling, Pilar Rivera Gil, Feng Zhang, Marco Zanella, and Wolfgang J Parak. Biological applications of gold nanoparticles. *Chemical Society reviews*, 37(9):1896, Jan 2008.
- [67] B Devika Chithrani and Warren C W Chan. Elucidating the mechanism of cellular uptake and removal of protein-coated gold nanoparticles of different sizes and shapes. *Nano Lett.*, 7(6):1542–50, Jun 2007.
- [68] B Devika Chithrani, Arezou A Ghazani, , and Warren C W Chan. Determining the size and shape dependence of gold nanoparticle uptake into mammalian cells. *Nano Letters*, 6:662 – 668, 2006.
- [69] T Kim, C.H Lee, S.W Joo, and K Lee. Kinetics of gold nanoparticle aggregation: experiments and modeling. *Journal of Colloid And Interface Science*, 318(2):238–243, 2008.
- [70] Yu Pan, Annika Leifert, David Ruau, Sabine Neuss, Jörg Bornemann, Günter Schmid, Wolfgang Brandau, Ulrich Simon, and Willi Jahnen-Dechent. Gold nanoparticles of diameter 1.4ÅÅnm trigger necrosis by oxidative stress and mitochondrial damage. *Small*, 5(18):2067–2076, Sep 2009.

- [71] Yu Pan, Sabine Neuss, Annika Leifert, Monika Fischler, Fei Wen, Ulrich Simon, Günter Schmid, Wolfgang Brandau, and Willi Jahnen-Dechent. Size-dependent cytotoxicity of gold nanoparticles. *Small*, 3(11):1941–1949, Nov 2007.
- [72] Alexandre Albanese and Warren C W Chan. Effect of gold nanoparticle aggregation on cell uptake and toxicity. *ACS Nano*, 5:5478–5489, 2011.
- [73] Junko Okuda-Shimazaki, Saiko Takaku, Koki Kanehira, Shunji Sonezaki, and Akiyohshi Taniguchi. Effects of titanium dioxide nanoparticle aggregate size on gene expression. *Int J Mol Sci*, 11(6):2383–92, Jan 2010.
- [74] Takami Akagi, Fumiaki Shima, and Mitsuru Akashi. Intracellular degradation and distribution of protein-encapsulated amphiphilic poly(amino acid) nanoparticles. *Biomaterials*, 32(21):4959–4967, Jul 2011.
- [75] Alaaldin M Alkilany and Catherine J Murphy. Toxicity and cellular uptake of gold nanoparticles: what we have learned so far? *J Nanopart Res*, 12(7):2313–2333, Sep 2010.
- [76] B Talcott and M S Moore. Getting across the nuclear pore complex. *Trends in Cell Biology*, 9(8):312–8, Aug 1999.
- [77] H Akita, R Ito, I A Khalil, S Futaki, and H Harashima. Quantitative three-dimensional analysis of the intracellular trafficking of plasmid dna transfected by a nonviral gene delivery system using confocal laser scanning microscopy. *Mol Ther*, 9(3):443–51, Mar 2004.
- [78] Boerries Brandenburg and Xiaowei Zhuang. Virus trafficking – learning from single-virus tracking. *Nat Rev Microbiol*, 5(3):197–208, Mar 2007.
- [79] Hunter H Chen, Yi-Ping Ho, Xuan Jiang, Hai-Quan Mao, Tza-Huei Wang, and Kam W Leong. Quantitative comparison of intracellular unpacking kinetics of polyplexes by a model constructed from quantum dot-fret. *Molecular Therapy*, 16(2):324–332, Feb 2008.
- [80] Susumu Hama, Hidetaka Akita, Rie Ito, Hiroyuki Mizuguchi, Takao Hayakawa, and Hideyoshi Harashima. Quantitative comparison of intracellular trafficking and nuclear transcription between adenoviral and lipoplex systems. *Mol Ther*, 13(4):786–94, Apr 2006.
- [81] M. J Rust, M Lakadamyali, B Brandenburg, and X Zhuang. Single-virus tracking in live cells. *Cold Spring Harbor Protocols*, 2011(9):pdb.top065623–pdb.top065623, Sep 2011.
- [82] Johann Kecht, Axel Schlossbauer, and Thomas Bein. Selective functionalization of the outer and inner surfaces in mesoporous silica nanoparticles. *Chem Mater*, 20:7207–7214, 2008.

- [83] B Dubertret. In vivo imaging of quantum dots encapsulated in phospholipid micelles. *Science*, 298(5599):1759–1762, Nov 2002.
- [84] Sung-Hee Park, Seong-Geun Oh, Ji-Young Mun, and Sung-Sik Han. Loading of gold nanoparticles inside the dppc bilayers of liposome and their effects on membrane fluidities. *Colloids Surf B Biointerfaces*, 48(2):112–8, Mar 2006.
- [85] Andreas O Hohner, Maria Pamela C David, and Joachim O Rädler. Controlled solvent-exchange deposition of phospholipid membranes onto solid surfaces. *Biointerphases*, 5(1):1–8, Mar 2010.
- [86] Valentina Cauda, Hanna Engelke, Anna Sauer, Delphine Arcizet, Joachim Rädler, and Thomas Bein. Colchicine-loaded lipid bilayer-coated 50 nm mesoporous nanoparticles efficiently induce microtubule depolymerization upon cell uptake. *Nano Lett.*, 10(7):2484–2492, Jul 2010.
- [87] Simon Youssef, Sebastian Gude, and Joachim O. Radler. Automated tracking in live-cell time-lapse movies. *Integr. Biol.*, 3:1095–1101, 2011.
- [88] Guangpu Li and Hong Qian. Kinetic timing: a novel mechanism that improves the accuracy of gtpase timers in endosome fusion and other biological processes. *Traffic*, 3(4):249–55, Apr 2002.
- [89] C Greulich, J Diendorf, J Geßmann, T Simon, T Habijan, G Eggeler, T A Schildhauer, M Epple, and M Köller. Cell type-specific responses of peripheral blood mononuclear cells to silver nanoparticles. *Acta Biomaterialia*, 7(9):3505–3514, Sep 2011.
- [90] Yvonne Williams, Alyona Sukhanova, Małgorzata Nowostawska, Anthony M Davies, Siobhan Mitchell, Vladimir Oleinikov, Yurii Gun’ko, Igor Nabiev, Dermot Kelleher, and Yuri Volkov. Probing cell-type-specific intracellular nanoscale barriers using size-tuned quantum dots. *Small*, 5(22):2581–2588, Nov 2009.
- [91] Zhu-Nan Li, Byeong-Jae Lee, William A Langley, Konrad C Bradley, Rupert J Russell, and David A Steinhauer. Length requirements for membrane fusion of influenza virus hemagglutinin peptide linkers to transmembrane or fusion peptide domains. *Journal of Virology*, 82(13):6337–48, Jul 2008.
- [92] Aurélien Lorin, Laurence Lins, Vincent Stroobant, Robert Brasseur, and Benoit Charletoeux. Determination of the minimal fusion peptide of bovine leukemia virus gp30. *Biochem Biophys Res Commun*, 355(3):649–53, Apr 2007.
- [93] S Fawell, J Seery, Y Daikh, C Moore, L L Chen, B Pepinsky, and J Barsoum. Tat-mediated delivery of heterologous proteins into cells. *Proc Natl Acad Sci USA*, 91(2):664–8, Jan 1994.

- [94] E Vivès, P Brodin, and B Lebleu. A truncated hiv-1 tat protein basic domain rapidly translocates through the plasma membrane and accumulates in the cell nucleus. *J Biol Chem*, 272(25):16010–7, Jun 1997.
- [95] Jehangir S Wadia, Radu V Stan, and Steven F Dowdy. Transducible tat-ha fusogenic peptide enhances escape of tat-fusion proteins after lipid raft macropinocytosis. *Nat Med*, 10(3):310–315, Mar 2004.
- [96] Iris Geisler and Jean Chmielewski. Probing length effects and mechanism of cell penetrating agents mounted on a polyproline helix scaffold. *Bioorg Med Chem Lett*, 17(10):2765–8, May 2007.
- [97] William A Langley, Sudha Thoennes, Konrad C Bradley, Summer E Galloway, Ganesh R Talekar, Sandra F Cummings, Eva Varecková, Rupert J Russell, and David A Steinhauer. Single residue deletions along the length of the influenza ha fusion peptide lead to inhibition of membrane fusion function. *Virology*, 394(2):321–330, Nov 2009.
- [98] Ira Pastan, Raffit Hassan, David J Fitzgerald, and Robert J Kreitman. Immunotoxin therapy of cancer. *Nat Rev Cancer*, 6(7):559–65, Jul 2006.
- [99] M.P Dobay, A Dobay, J Bantang, and E Mendoza. How many trimers? modeling influenza virus fusion yields a minimum aggregate size of six trimers, three of which are fusogenic. *Mol. BioSyst.*, 7(10):2741–2749, 2011.
- [100] M Kwiatkowska, G Norman, and D Parker. Prism 4.0: Verification of probabilistic real-time systems. *In Proc. 23rd International Conference on Computer Aided Verification (CAV'11) LNCS*, 6806:585–591, 2011.
- [101] L Lins and R Brasseur. Tilted peptides: a structural motif involved in protein membrane insertion? *Journal of Peptide Science*, 14(4):416–422, 2008.
- [102] K S Rostand and J D Esko. Microbial adherence to and invasion through proteoglycans. *Infection and Immunity*, 65(1):1–8, Jan 1997.
- [103] T Surrey and F Jähnig. Refolding and oriented insertion of a membrane protein into a lipid bilayer. *Proc Natl Acad Sci USA*, 89(16):7457–61, Aug 1992.
- [104] Michèle A Barocchi, Vega Massignani, and Rino Rappuoli. Opinion: Cell entry machines: a common theme in nature? *Nat Rev Microbiol*, 3(4):349–58, Apr 2005.
- [105] R Weissleder, K Kelly, EY Sun, T Shtatland, and L Josephson. Cell-specific targeting of nanoparticles by multivalent attachment of small molecules. *Nat Biotechnol*, 23(11):1418–1423, 2005.

- 
- [106] D C Drummond, O Meyer, K Hong, D B Kirpotin, and D Papahadjopoulos. Optimizing liposomes for delivery of chemotherapeutic agents to solid tumors. *Pharmacol Rev*, 51(4):691–743, Dec 1999.
- [107] Ann-Sofie Cans, Nathan Wittenberg, Roger Karlsson, Leslie Sombers, Mattias Karlsson, Owe Orwar, and Andrew Ewing. Artificial cells: unique insights into exocytosis using liposomes and lipid nanotubes. *Proc Natl Acad Sci USA*, 100(2):400–4, Jan 2003.
- [108] Nicholas A W Bell, Christian R Engst, Marc Ablay, Giorgio Divitini, Caterina Ducati, Tim Liedl, and Ulrich F Keyser. Dna origami nanopores. *Nano Lett.*, 12(1):512–7, Jan 2012.
- [109] Carlos Ernesto Castro, Fabian Kilchherr, Do-Nyun Kim, Enrique Lin Shiao, Tobias Wauer, Philipp Wortmann, Mark Bathe, and Hendrik Dietz. A primer to scaffolded dna origami. *Nat Meth*, 8(3):221–229, Mar 2011.
- [110] Anton Kuzyk, Kimmo T Laitinen, and Päivi Törmä. Dna origami as a nanoscale template for protein assembly. *Nanotechnology*, 20(23):235305, Jun 2009.

# Acknowledgements

“At times our own light goes out and is rekindled by a spark from another person. Each of us has cause to think with deep gratitude of those who have lighted the flame within us.”

-Albert Schweitzer

The journey of the past three and a half years was the most fulfilling that I have taken so far – but expectedly, also the most difficult. I wish to thank wonderful fellow travelers who have given me a home away from home, who have provided guidance, and who have rekindled my light.

I first wish to thank my professor, **Dr. Joachim Rädler**, who gave me the chance to work on this interesting project; who challenged me to look at phenomena from a different perspective than what I came with; who provided a possibility for me to perform both computational and experimental work; and who gave me the freedom to explore my way through the project – with the occasional reminder to make sure that I always remain on track.

**Dr. Eduardo Mendoza** for the constant guidance that has allowed me to outperform what I thought was my best, and for always taking the time to provide constructive feedback – even at 4 A.M! I am especially grateful that he arranged my first research visit to Munich, which was pivotal in both my professional and personal life.

The **Deutscher Akademischer Austausch Dienst (DAAD)**, and **Frau Elke Burbach**, in particular, for financing my studies in Germany – a privilege of a lifetime.

**Prof. Thomas Bein**, for allowing me to continue on the very exciting nanoparticle project, and for inviting me to his group meetings, where I could get a glimpse of the nitty-gritty behind a very exciting field. Many thanks, in particular, to **Alexandra Schmidt**, without whom all the endosome lysis work would not have been possible.

The **Center for NanoScience (CeNS)**, for financially supporting the endosome lysis project.

**Alicia Piera Alberola**, for patiently re-honing my skills in my microscopy and cell culture, for always keeping **purplesheep** running, and for the friendly hallway discussions. Thank you as well for your patience with the scripts sent your way!

**Johnrob Bantang**, for all the friendly after-coffee discussions on physics and the occasional debates on philosophy and PABM – which have gotten us both thinking at one point or the other.

**Hanna Engelke**, for helping me optimize setups for the nanoparticle experiments, and for answering some naive first questions on these.

**Tobias Stögbauer**, for getting me started on all things lipid.

**Gerlinde Schwake** and **Susi Kempter**, my first roommates in 2005, who have been the backbone of not just my labwork – but of everyone else’s.

Special thanks to **Carolin Leonhardt** for transforming my German version of the summary from atrocious to elegant, and for having the patience for that bit of pi (calc).

**Margarete Meixner**, for making my life easier in a lot of ways – from arranging meetings to making calls that saved my salary; for all the very interesting conversations over late lunch; and for bearing with my not-always-intelligible German conversations.

The **Filipino community in Munich** for the fiesta atmosphere, for the occasional death-daring outdoor adventures, and for being a constant reminder of the best things back home. Special thanks to **Florence-Balagtas Fernandez** for taking the time for gastronomic and retail adventures when stress levels ran particularly high...

My former mentors, **Profs. Gisela P. Concepcion, Edurado A. Padlan, Evangel P. Quiwa**, and **Joyce S. Ibana**, from whom I acquired the discipline and tenacity required in endeavors such as this.

My heartfelt thanks to **Dr. Thomas Ligon** and **Anneliese Ligon-Klein**, for graciously hosting me in my last visits to Munich, for all the interesting discussions that range from science to life (and all the interesting bits in between) over dinner with either silky or persistent tannins, for the lovely walks in the countryside, and for the warm friendship.

And to **Akos**, for the unconditional support, the wonderful music – and for answering that occasional scientific question at 3 A.M., when I couldn’t bug anyone else. Köszönöm szépen!

# Lebenslauf

## Personal data

Maria Pamela Dobay, geb. David  
Schulhausstrasse 40A  
Zurich, Schweiz 8002  
born 19 January 1982, Quezon City, Philippines

## Education

- 2008 - 2012 PhD student, Physics, in the group of Prof. Dr. J. O. Raedler,  
Ludwig-Maximilians-Universitaet, Muenchen  
(DAAD scholar)
- 2005 - 2008 Master of Science, Computer Science  
University of the Philippines, Diliman
- 1999 - 2003 Bachelor of Science, Molecular Biology and Biotechnology  
(DOST scholar)  
University of the Philippines, Diliman
- 1995 - 1999 Philippine Science High School, Diliman, Quezon City

## Experience

- 2007 - 2008 Instructor, Department of Computer Science  
University of the Philippines, Diliman
- 2006 - 2008 Developer, EUCLOCK information system,  
EUCLOCK project
- 2005 Visiting researcher, Department of Physics  
LS Prof. Dr. J. O. Raedler,  
Ludwig-Maximilians-Universitaet, Muenchen
- 2003 - 2005 Researcher, Marine Science Institute  
University of the Philippines, Diliman, Quezon City