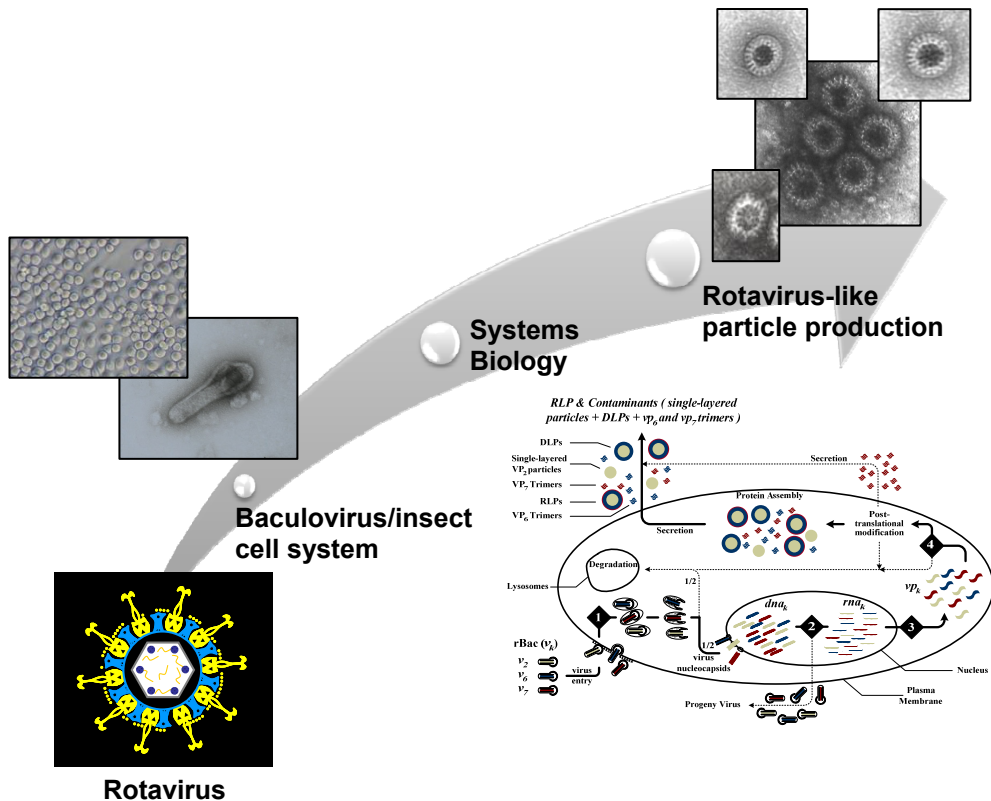


# PRODUCTION OPTIMIZATION OF ROTAVIRUS-LIKE PARTICLES

## A SYSTEMS BIOLOGY APPROACH



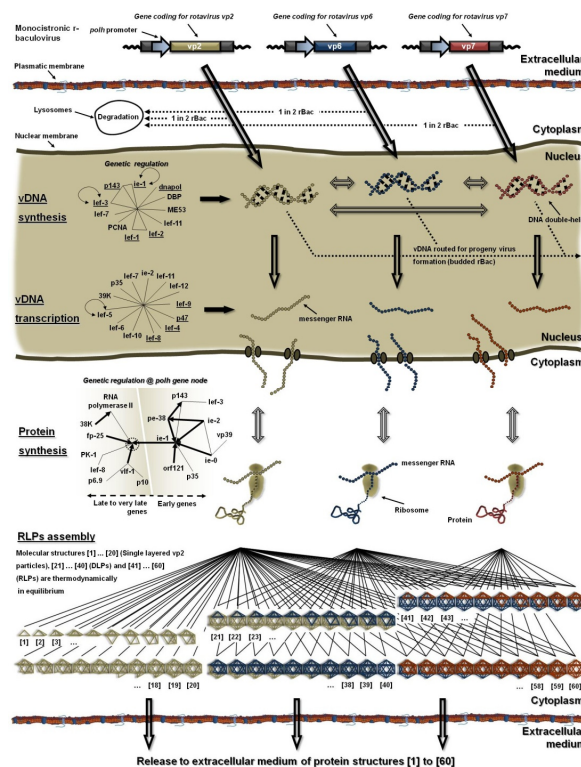
António Manuel Missionário Roldão



Dissertation presented to obtain a Ph.D. degree in Engineering and Technology Sciences, Systems Biology at the Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa

Oeiras, June 2010

To understand the entirety of the processes happening during the production of a rotavirus-like particle, from the genetics of the viral vector to the infection strategy, titer and structure of assembled particles, a bottom-up Systems Biology approach is required.



This thesis provides a set of computational models and algorithms which represent a step forward to the *in silico* implementation of the baculovirus/insect cells system. Through simulation and numerical optimization codes, novel and improved optimization schemes for production of virus particles, viral proteins and/or complex protein structures such as virus-like particles can be developed.

# PRODUCTION OPTIMIZATION OF ROTAVIRUS- LIKE PARTICLES

*A SYSTEMS BIOLOGY APPROACH*

António Manuel Missionário Roldão

*Dissertation presented to obtain a Ph.D. degree in Engineering and  
Technology Sciences, Systems Biology at the Instituto de Tecnologia  
Química e Biológica, Universidade Nova de Lisboa*

*Supervisors:*

Rui M. F. Oliveira and Manuel J. T. Carrondo



Instituto de Tecnologia Química e Biológica

Universidade Nova de Lisboa

Oeiras, June 2010

***Production Optimization of Rotavirus-Like Particles - A Systems  
Biology Approach***

by António Roldão

First edition: June 2010

Copyright number:

*Cover images:*

*Front:* schematic representation of the rotavirus-like particle optimization process; illustration of a rotavirus, electron micrograph of negatively stained baculovirus and rotavirus-like particles, network of heterologous protein expression using the baculovirus/insect cell system

*Back:* genetic/molecular regulatory network of rotavirus-like particle production in the baculovirus/insect cell system

© 2010

António Manuel Missionário Roldão

Ph.D. Thesis

Animal Cell Technology Unit

Instituto de Tecnologia Química e Biológica – Universidade Nova de Lisboa/

Instituto de Biologia Experimental e Tecnológica

Av. da República, Apt. 127, 2781-901 Oeiras, Portugal

<http://www.itqb.unl.pt>

<http://www.ibet.pt>



## ***Supervisors***

Dr. Rui Manuel Freitas Oliveira, Auxiliary Professor at Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Caparica, Portugal.

Dr. Manuel José Teixeira Carrondo, Cathedratric Professor of Chemical and Biochemical Engineering at Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Caparica, Portugal, and CEO of IBET, Oeiras, Portugal.

## ***Jury***

Dr. Bruce Tidor, Professor of Biological Engineering and Computer Science at Massachusetts Institute of Technology, Massachusetts, USA.

Dr. Eugénio Campos Ferreira, Associate Professor at Biological Engineering Department, Universidade do Minho, Braga, Portugal.

Dr. Vicente Bernal Sanchez, Post-Doctoral Research at Faculdade de Química, Universidade de Murcia, Murcia, Spain.

Dr. João Carlos Lima, Auxiliary Professor, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Caparica, Portugal.

Dr. Paula Maria Marques Alves, Principal Investigator and Head of the Animal Cell Technology Laboratory at ITQB and Executive Director of IBET, Oeiras, Portugal.

## **Foreword**

The present dissertation is the result of four years of research at the Animal Cell Technology Unit of the Instituto de Tecnologia Química e Biológica-Universidade Nova de Lisboa/Instituto de Biologia Experimental e Tecnológica, under the supervision of Dr. Rui Oliveira and Prof. Manuel Carrondo.

This thesis intends to provide a set of computational models and algorithms, which represent a step forward to the *in silico* implementation of the baculovirus/insect cells system, to assist the design of novel and improved optimization schemes for the production of virus particles, viral proteins and/or complex protein structures such as virus-like particles.

Aos meus pais, irmã e avós

À Madalena e Francisco



## **Acknowledgements**

I would like to acknowledge all the people directly or indirectly involved in this thesis, which supported and helped me during this long but rewarding journey.

To my “Great” supervisor, Professor Manuel Carrondo, I would like to express my sincere gratitude for the opportunity given to perform the work presented in this thesis. Its outstanding personality, life experience and thoughtful advices have made me grow as a person and as a scientist. You are one of a kind; my deepest respect. It was an enormous pleasure working with you Professor. I sincerely hope that we could meet again in the near future.

To my “Systems Biology” supervisor, Dr. Rui Oliveira: I owe most of this thesis to you. You are, far-off, the brightest professor/scientist that one could expect to have when trying to bring mathematical formulas to the complex field of biology. I have learnt a lot with you, especially from our long discussions in Caparica where we stood up hours in a row trying to discover how to describe mathematically a specific feature of our system. For all the hard work developed, my gratitude.

The “Boss/Friend” Dr. Paula Alves: Personally, I consider you a role model. You are an entrepreneur with a clear vision on what science should be, a great scientist, an intelligent group leader and a good friend. I could have not asked for a better boss. You gave me the opportunity to work in the Animal Cell Technology Unit, a group of excellence to which I am truly honoured to belong. This thesis is also yours.

To the former “Rotavirus team” for their priceless contribution to this thesis and helpful discussions, namely Dr. Helena Vieira, Dr. Candida Mellado, M.Sc. Catarina Estevão and Eng. Ana Luísa Escapa.

A special acknowledgment should be made to Dr. Jean Cohen from the CNRS-INRA, the main responsible for the initiation of this project. To other partners of the project for providing the recombinant baculoviruses, namely Dr. Didier Poncet and Dr. Annie Charpilienne from CNRS-INRA and Dr. Polly Roy from the LSHTM. To all the people indirectly involved in the *in vitro* experiments, namely Dr. Ana Luisa Simplicio (IBET) and Dr. Max Ciarlet (Merck & Co., Inc.).

To Dr Pedro Cruz, Dr. Catarina Brito, Dr. Vicente Bernal, Dr. Ana Teixeira, Eng. Tiago Vicente and Nuno Carinhas for their excellent advices, fruitful and interesting discussions and valuable comments. Bright minds with enormous curiosity for the unknown and always receptive to new and challenging ideas.

I am grateful to Prof João C. Lima from the FCT-UNL for introducing me to the thermodynamic field and from whom I have learned a lot on protein and molecular interactions.

To all the present members of the Animal Cell Technology Unit for creating a competitive but healthy working environment, especially the ones that accompanied me since my infancy times in the lab: Cristina Peixoto, Ana Sofia Coroadinha, Rita Malpique and Sofia Leite. To former members of the Animal Cell Technology Unit Tiago Ferreira and Isabel Marcelino, dear friends and great young scientists, to Ana Mendes, José Bragança and Monica Martinez.

I am deeply grateful to Guida, Marcos, Carina, Ricardo Perdigão e Teresa for all the support, loyalty, encouragement, laugh and friendship. Wherever I will go, I will always have you in my heart. Many thanks for being with me and for being as you are.

To the entities that funded the research: Fundação para a Ciência e Tecnologia (FCT) and Fundo Social Europeu (FSE) (SFRH/BD/21910/2005 and POCTI/BIO/55975/2004) and from the European Commission project (QLRT-2001-01249 and Baculogenes, LSHB-2006-037541).

Aos meus grandes amigos de sempre Susana, Luís, Sara, Helder e Íris por todo o apoio e amizade. Aos meus grandes colegas de futebol Pinguço, Viseu, Caeiro, Batatoon, Gil, Telmo, Bruno, Rui e muitos outros que tanto contribuíram para que libertasse todo o stress acumulado ao longo de um dia de trabalho, vai o meu bem hajam.

Não poderia terminar sem agradecer às pessoas a quem devo tudo na vida, os meus pais. Sem vocês não seria o que sou hoje. À minha compreensiva e adorável irmã, à minha querida avó Joaquina, ao grande amigo Fernando e aos meus sobrinhos Francisco e Madalena, pequeninos pedaços de alegria que fazem as minhas delicias.

*“The secret of getting ahead is getting started. The secret of getting started is breaking your complex, overwhelming tasks into small manageable tasks, and then starting on the first one.”*

Mark Twain





## ***Thesis publications***

**António Roldão**, Maria Candida M. Mellado, J. C. Lima, Manuel J.T. Carrondo, Paula M. Alves and R. Oliveira, 2010. A Thermodynamic Equilibrium-Based Model for *in vitro* Analysis and Design of Virus-Like Particles (VLP) Assembly: The Case of Rotavirus VLP, *Biophysical Journal* (under review).

**António Roldão**, Rui Oliveira, Manuel J. T. Carrondo and Paula M. Alves, 2009. Error Assessment in Recombinant Baculovirus Titration: Evaluation of Different Methods. *Journal of Virological Methods*, 159(1): 69-80.

**António Roldão**, Manuel J.T. Carrondo, Paula M. Alves and Rui Oliveira, 2008. Stochastic simulation of protein expression in the baculovirus/insect cells system. *Computers and Chemical Engineering*, 32(1-2): 68-77.

**António Roldão**, Helena L.A. Vieira, Annie Charpilienne, Didier Poncet, Polly Roy, Manuel J.T. Carrondo, Paula M. Alves and Rui Oliveira, 2007. Modeling Rotavirus-Like Particles production in a baculovirus expression vector system: infection kinetics, baculovirus DNA replication, mRNA synthesis and protein production. *Journal of Biotechnology*, 128(4): 875-894.

**António Roldão**, Helena L. A. Vieira, Paula M. Alves, Rui Oliveira, Manuel J.T. Carrondo, 2006. Intracellular dynamics in Rotavirus-like particles production: Evaluation of multigene and monocistronic infection strategies. *Process Biochemistry* 41, 2188-2199.

Helena L.A. Vieira, Catarina Estêvão, **António Roldão**, Cristina C. Peixoto, Marcos F.Q. Sousa, Pedro E. Cruz, Manuel J.T. Carrondo and Paula M. Alves, 2005. Triple layered rotavirus VLP production: Kinetics of vector replication, mRNA stability and recombinant protein production. *Journal of Biotechnology*, 120 (1): 72-82.

## ***Publications in conference proceedings***

**António Roldão**, Helena L.A. Vieira, Manuel J.T. Carrondo, Paula M. Alves and

Rui Oliveira, 2005. Modeling and Control of Rotavirus Virus-Like Particle Production: Baculovirus DNA Replication, mRNA Synthesis and Protein Production. Abstracts of the 9<sup>th</sup> International Chemical Engineering Conference, Coimbra, Portugal 21<sup>st</sup> – 23<sup>rd</sup> September, 2005.

**António Roldão**, Helena L. A. Vieira, Manuel J.T. Carrondo, Paula M. Alves and R. Oliveira, 2006. Rotavirus-Like Particle Production: Simulation of Protein Production and Particle Assembly. *Computer Aided Chemical Engineering*, 21(Part 2): 1673-1678. Abstracts of the ESCAPE-16 + PSE, Garmisch-Partenkirchen, Germany, July 9<sup>th</sup> – 13<sup>rd</sup>, 2006.

**António Roldão**, Manuel J.T. Carrondo, Paula M. Alves and Rui Oliveira, 2008. Stochastic Modeling of the Baculovirus/Insect Cells System: Prediction of Recombinant Protein Expression. Abstracts of the 10<sup>th</sup> International Chemical Engineering Conference, Braga, Portugal 4<sup>th</sup> – 6<sup>th</sup> September, 2008.

**A. Roldão**, M.J.T. Carrondo, P.M. Alves, R. Oliveira, 2009. *In silico* rotavirus-like particle production in the recombinant baculovirus/insect cell system. *New Biotechnology*, 25(1): S356-S357. Abstracts of the 14<sup>th</sup> European Congress on Biotechnology, Barcelona, Spain 13<sup>th</sup> – 16<sup>th</sup> September, 2009.

## ***Book Chapter***

**António Roldão**, Ana Carina Silva, Maria Candida M. Mellado, Paula M. Alves, Manuel J. T. Carrondo, 2009. Viruses and virus-like particles in biotechnology: fundamentals and applications. In: Moo-Young and Butler Eds. *Comprehensive Biotechnology*, 2<sup>nd</sup> edition, *Volume 1: Scientific Fundamentals in Biotechnology*, Elsevier, Oxford (submitted).

## ***Dissertation Abstract***

Rotavirus-like particles (RLPs), a vaccine candidate against rotavirus disease, were produced by infecting *Spodoptera frugiperda* Sf-9 cells with genetically engineered recombinant baculoviruses. RLPs are spherically shaped particles composed by three viral proteins (vp) of rotavirus, vp2, vp6 and vp7, arranged in a triple layered structure.

A diversity of protein structures, other than the correctly assembled RLP, are observed at the end of a typical production run suggesting that the protein assembly process is rather inefficient. Contaminants such as trimers of vp6 and vp7, vp6 tube-like structures, single-layered vp2 particles, double layered particles of vp2 and vp6 or RLPs lacking one or more subunits represent almost 88% of the total mass of proteins expressed. Thus, optimal control of protein expression concomitant with efficient particle assembly are critical factors for economical RLP production in the baculovirus/insect cells system.

In this thesis, experimental activities and mathematical modelling are merged together towards the development of an *in silico* baculovirus/insect cells system implementation with the ability to simulate and predict multiple protein expression and posterior self assembly into virus-like particles (VLPs), using RLPs as model system.

In Chapter 1, the state of the art on existing mathematical models developed for the baculovirus/insect cell system is presented. Brief bioengineering notions involved in this expression system, types of mathematical models and VLPs are also provided.

Since the lack of accuracy of baculovirus titers directly affects the multiplicity of infection (MOI) and impacts negatively on batch consistency and process optimization, a comparative study was performed in Chapter 2

to identify the titration methods that better combines accuracy with the method's cost and work intensity. The best compromise was achieved with the endpoint dilution, microculture tetrazolium and flow cytometric assays. Nonetheless, checking titer's accuracy against a baculovirus reference stock or by titration using two different methods and verification of the variability of results is always advisable. Based on these data, it may be realistic to expect a MOI variability of around 20%.

In Chapter 3, the differences between deterministic and stochastic infection, and the implications of stochastic baculovirus infection on protein expression at high and low MOIs are discussed. The saturation in intracellular protein expression occurs due to the existence of a maximum translation capacity; hence, there is no advantage in increasing the MOI to very high levels as the translation machinery saturates, thus resulting in no improvement in protein expression. The total concentration of viral protein was shown to be differently affected by the MOI. At MOIs  $> 10 \text{ virus.cell}^{-1}$ , no significant improvement in protein concentration was achieved as a consequence of protein expression saturation. In addition, since the infection at such high MOIs is essentially synchronous and instantaneous, the deterministic and stochastic models converged to the same solution. At lower MOIs, protein concentration was maximized as result of higher infected cell concentration. The outputs of stochastic and deterministic model diverged completely since infection is asynchronous and ruled by stochasticity. Interestingly, the gene size did not affect the optimal MOI ranges for protein expression but impacted negatively on the expression levels: the higher the gene size, the lower is the expression of the viral. This means that different proteins are expected to have optimal productivities around the same MOI obtained for vp2: between 0.01 and  $0.1 \text{ virus.cell}^{-1}$ .

A thermodynamic equilibrium-based model describing the assembly of RLPs was proposed in Chapter 4. The model establishes a relationship between particle formation and the Gibbs free energy of protein subunit association ( $\Delta G^0_n$ ) and protein concentration. Model simulations have shown that there is a minimum concentration of protein structural subunits ensuring a measurable particle formation, i.e. a pseudo-critical concentration level. Importantly, this concentration threshold was found to be correlated with  $\Delta G^0_n$ : as  $\Delta G^0_n$  increases, the protein's pseudo-critical concentration also increases. The assembly and disassembly efficiencies of RLPs were found to be correlated with Gibbs free energy of vp7 subunit association ( $\Delta G^0_{RLP}$ ): as  $\Delta G^0_{RLP}$  increases, the assembly increases and the disassembly decreases accordingly. Interestingly, this dependency is not linear. For Gibbs free energy values below  $-2 \text{ kcal.mol}^{-1}$ , changes in Gibbs free energies do not affect significantly the assembly of RLPs. Above this threshold, a crisp transition from intact icosahedra to unassembled subunits is observed. Noteworthy was the impact of Gibbs free energy on optimal initial proteins concentration. For low Gibbs free energies, thermodynamics maximizes the formation of intact icosahedra (RLP) whenever proteins are provided at the exact RLP stoichiometric ratio. At extremely high Gibbs free energies, there is a translocation of the proteins pseudo-critical concentrations; so, optimum concentrations were no longer coincident with the stoichiometry of the particle. Since host cells can hardly be manipulated to operate in such a narrow region of Gibbs free energies and protein concentrations, the formation of malformed particles is likely to be inevitable *in vivo*.

In Chapter 5, a deterministic structured model was developed to understand and quantify the effect of different infection strategies on RLP production. Two major outcomes of the model were the vDNA replication rate, constant at  $0.19 \pm 0.01 \text{ h}^{-1}$  irrespective of the baculovirus vector

(monocistronic or tricistronic), and the apparent higher potential of *p10* promoter for vp7 expression in comparison to the *polh* promoter. Analysis of the time windows for vDNA, mRNA and protein synthesis revealed that these intracellular processes are highly correlated and MOI dependent. Indeed, higher intracellular vDNA concentrations induce higher mRNA transcription levels and, ultimately, higher intracellular protein concentrations. In addition, the higher the MOI, the more severe is the “metabolic burden” and, consequently, the lower the time window for intracellular vDNA, mRNA and protein synthesis. With the use of tricistronic baculoviruses (single-infection), the possibility of infected cells not carrying all genes required for RLP assembly (genes coding for vp2, vp6 and vp7), as it happens with co-infection experiments, was overcome. Therefore, protein limitations were remarkably less severe, which, in the end, led to the formation of more RLPs when compared to co-infection.

An *in silico* analysis of the biological system was performed, namely concerning the sensitivity of RLPs and constitutive proteins concentration in relation to model parameters (Chapter 6) and to process degrees of freedom such as the MOI or  $\Delta G^0_n$  (Chapters 3 to 6). Model outputs were found to be more sensitive to changes in parameters related with vDNA and protein synthesis, such as  $\delta_{10}$ ,  $k_{\text{RDNA}}$ ,  $S_2$ ,  $mp_2$  and  $k_{\text{rib,elong}}$ , than to those linked to cell growth, infection process, virus budding, re-infection and mRNA.

In Chapter 6, the theoretical effect of MOI and promoter strength on RLP production was investigated for single and co-infection strategies using the stochastic model developed in Chapter 3. At single-infections, RLP production was optimized by manipulating the MOI (lower MOIs) and by redesigning baculovirus vectors, in which the promoters controlling viral protein expression can be selected according to their strength (lowering the strength of the promoter controlling vp7 expression in relation to the other

two promoters). For co-infection strategies, high MOIs maximized RLP production; the redesign of baculovirus vectors configured a somewhat modest improvement potential. Overall, single-infection strategies seem to offer significant advantages over co-infection experiments. This is clearly evident at low MOIs where the higher cell density achieved at single-infection led to higher RLP production (30-60 mg.L<sup>-1</sup> against the 6-12 mg.L<sup>-1</sup> at co-infection). The assembly efficiencies were also higher: 72% against 20-40% at co-infection. At high MOIs, the difference in model outputs (assembly efficiency and intracellular and total RLP concentration) between both strategies was not significant.

Chapter 7 discusses the implications of the findings presented in the other chapters. It summarizes the main achievements of this thesis, identifies the main pitfalls of the computational tools developed in previous chapters for the optimisation of the baculovirus/insect cell system and outlines potential strategies for the production of RLPs in a more efficient way.

Summarizing, this thesis provides a set of computational models and algorithms which represent a step forward to the *in silico* implementation of the baculovirus/insect cells system. Through simulation and numerical optimization codes, novel and improved optimization schemes for production of virus particles, viral proteins and/or complex protein structures such as VLPs can be developed.





## ***Resumo da Dissertação***

As partículas semelhantes ao rotavírus (RLPs), vacinas candidatas contra a doença do rotavírus, são normalmente produzidas ao infectar células de insecto (*Spodoptera frugiperda* Sf-9) com baculovírus recombinantes geneticamente modificados. Estas partículas de formato esférico são compostas por três proteínas virais (vp) do rotavírus, nomeadamente vp2, vp6 e vp7, dispostas numa estrutura de tripla camada.

A diversidade de estruturas proteicas observada no final de um ciclo de produção de RLPs indica que o processo de montagem é ineficiente. Trímeros de vp6 e vp7, estruturas tubulares de vp6, partículas esféricas compostas por vp2, partículas esféricas de dupla camada compostas por vp2 e vp6 (DLP) ou RLPs que têm uma ou mais subunidades estruturais em falta são exemplos de contaminantes do processo, representando quase 88% da massa total de proteínas expressas. Como tal, o controlo da expressão das proteínas juntamente com a montagem eficiente da partícula são factores essenciais para uma produção económica de RLPs no sistema de baculovírus/células de insecto.

Nesta tese, actividades experimentais e modelação matemática são combinadas com o objectivo de desenvolver e implementar um sistema de baculovírus/células de insecto virtual com a capacidade de simular e prever a expressão de múltiplas proteínas e a sua posterior montagem em partículas semelhantes a vírus (VLPs), usando as RLP como modelo.

No Capítulo 1, são apresentados os modelos matemáticos desenvolvidos para o sistema de baculovírus/células de insecto. Breves noções sobre as questões de bioengenharia deste sistema de expressão, as diferentes categorias de modelos matemáticos e os tipos de VLPs existentes são igualmente apresentadas.

Uma vez que a imprecisão nos títulos dos baculovírus afecta directamente a multiplicidade de infecção (MOI – número de vírus por célula) e tem um impacto negativo sobre a reprodutibilidade e optimização do processo de produção de RLPs, foi realizado um estudo comparativo de modo a identificar os métodos de titulação que melhor combinam precisão, custo do método e intensidade de trabalho (Capítulo 2). O melhor compromisso foi alcançado para os métodos “endpoint dilution”, “microculture tetrazolium” e “flow cytometry”. No entanto, é sempre aconselhável verificar a precisão do título estimado comparando-o com um “stock” de baculovírus de referência ou titulando o “stock” em questão por dois métodos diferentes e analisar a variabilidade dos resultados obtidos. Tendo em conta os resultados obtidos, é realístico esperar que a MOI tenha uma variabilidade a rondar os 20%.

No Capítulo 3, foram discutidas as diferenças entre a infecção determinística e estocástica bem como as implicações da infecção estocástica na expressão de uma proteína a MOIs altas e baixas. Observou-se uma saturação na expressão de proteína, o que é justificado pela existência de uma capacidade máxima de tradução; como tal, não há vantagem em aumentar a MOI para valores muito elevados uma vez que a maquinaria de tradução satura, não havendo qualquer aumento na expressão da proteína. Foi verificado que a concentração total de proteína é afectada de diferente forma pela MOI. A MOIs  $> 10 \text{ virus.cell}^{-1}$ , não houve aumento significativo na concentração da proteína devido à saturação na expressão de proteína. Para além disso, uma vez que a infecção a MOIs muito elevadas é essencialmente síncrona e instantânea, o modelo de infecção determinístico e estocástico convergiram para a mesma solução. Para MOIs mais baixas, a concentração de proteína foi maximizada em resultado de uma maior concentração de células infectadas. Os modelos de infecção estocástico e determinístico divergiram uma vez que a

infecção é assíncrona e regulada pela estocasticidade. Interessantemente, o tamanho do gene não afectou o intervalo de MOIs ideal para a expressão de proteína, mas teve um impacto negativo sobre os níveis de expressão: quanto maior o tamanho do gene, menor é a expressão da proteína. Isto significa que é esperado que diferentes proteínas tenham produtividades óptimas à volta da mesma MOI que para a vp2: entre 0.01 e 0.1 virus.cell<sup>-1</sup>.

No Capítulo 4 foi desenvolvido um modelo matemático que descreve a montagem, passo a passo, de uma RLP e que se baseia no equilíbrio termodinâmico entre as diferentes estruturas moleculares que constituem a RLP. O modelo estabelece uma relação entre a formação de uma partícula e a energia livre de Gibbs ( $\Delta G_n^0$ ) e a concentração de proteína. As simulações do modelo mostraram que existe uma concentração mínima de subunidades estruturais de proteína que garante uma mensurável formação de partículas, isto é, uma concentração pseudo-crítica. Foi verificado que esta concentração crítica se encontra correlacionada com  $\Delta G_n^0$ : um aumento de  $\Delta G_n^0$  conduz a um aumento na concentração pseudo-crítica de proteína. Foi igualmente observado que a eficiência da montagem e desmontagem de uma RLP se encontra correlacionada com a energia livre de Gibbs de associação de subunidades de vp7 ( $\Delta G_{RLP}^0$ ): à medida que aumenta  $\Delta G_{RLP}^0$ , a eficiência da montagem aumenta e a eficiência da desmontagem diminui em conformidade. Curiosamente, esta dependência não é linear. Para valores de  $\Delta G_n^0$  menores que -2 kcal.mol<sup>-1</sup>, variações na energia livre de Gibbs não afectam significativamente a eficiência da montagem de RLP. Acima deste limiar, é observada uma nítida transição de estruturas icosaédricas intactas (RLP) para subunidades estruturais de proteína não associadas. A energia livre de Gibbs teve um impacto significativo na concentração inicial óptima de proteínas. Para energias livre de Gibbs baixas, a termodinâmica maximiza a formação de RLPs sempre que as proteínas sejam fornecidas na relação

estequiométrica de RLP exacta. Para energias livre de Gibbs extremamente altas, há uma translocação da concentração pseudo-crítica de proteínas; deste modo, a concentração ideal já não será coincidente com a relação estequiométrica de RLP. Uma vez que a célula hospedeira dificilmente pode ser manipulada de modo a operar numa região muito restrita de energias livre de Gibbs e de concentrações proteicas, a formação de partículas mal-formadas é muito provável que seja inevitável *in vivo*.

No Capítulo 5, um modelo determinístico/estruturado foi desenvolvido para compreender e quantificar o efeito de diferentes estratégias de infecção na produção de RLP. Dois resultados importantes foram extraídos do modelo: verificou-se que a taxa de replicação de vDNA ( $0.19 \pm 0.01 \text{ h}^{-1}$ ) é independente do vector viral (baculovírus monocistrónico ou tricistrónico) e que o promotor *p10* apresenta um maior potencial para a expressão de vp7 do que o promotor *polh*. Ao analisar os intervalos temporais correspondentes à síntese de vDNA, mRNA e proteína constatou-se que estes três processos intracelulares estão correlacionados e são dependentes da MOI. De facto, concentrações intracelulares elevadas de vDNA induzem níveis elevados de mRNA e de proteína. Para além disso, quanto maior for a MOI, mais severa é a sobrecarga metabólica e consequentemente menor é o intervalo de tempo para a síntese de vDNA, mRNA e proteína. A possibilidade de células infectadas não transportarem todos os genes necessários para a montagem da RLP (genes que codificam para vp2, vp6 e vp7), como acontece na co-infecção, foi ultrapassada usando baculovírus tricistrónicos (“single-infection”). Deste modo, as concentrações das três proteínas tornaram-se mais adequadas para a montagem da RLP, o que, no final, levou à formação de maior quantidade de RLP quando comparado com a co-infecção.

Foi realizada uma análise ao sistema biológico, nomeadamente sobre a sensibilidade da concentração de RLP e proteínas constituintes em relação aos parâmetros do modelo (Capítulo 6) e aos graus de liberdade do processo tais como a MOI e  $\Delta G^0_n$  (Capítulos 3 a 6). Os resultados do modelo mostraram-se mais sensíveis a variações nos parâmetros relacionados com a síntese de vDNA e de proteína ( $\delta_{10}$ ,  $k_{\text{RDNA}}$ ,  $S_2$ ,  $mp_2$  e  $k_{\text{rib,elong}}$ ) do que aos ligados ao crescimento celular, processo de infecção, libertação de novos baculovírus, re-infecção e mRNA.

No Capítulo 6, os efeitos teóricos da MOI e da força do promotor na produção de RLP foram investigados para as estratégias de “single-infection” e co-infecção usando, para tal, o modelo estocástico desenvolvido no Capítulo 3. No caso da estratégia “single-infection”, a produção de RLP foi otimizada através da manipulação da MOI (MOIs baixas) e do redesenho dos vectores virais, nos quais os promotores que controlam a expressão das proteínas virais podem ser seleccionados de acordo com a sua força (diminuindo a força do promotor que controla a expressão de vp7 em relação aos promotores que regulam a expressão de vp2 e vp6). Na estratégia de co-infecção, MOIs elevadas maximizaram a produção de RLP; o redesenho dos vectores virais configura-se uma alternativa com baixo potencial de melhoramento. Globalmente, a estratégia “single-infection” parece oferecer vantagens significativas sobre as experiências de co-infecção. Isto é claramente evidente para baixas MOIs onde a maior densidade celular alcançada na estratégia “single-infection” conduziu a uma maior produção de RLP (30-60 mg.L<sup>-1</sup> contra os 6-12 mg.L<sup>-1</sup> da co-infecção). As eficiências de montagem de RLPs também foram maiores: 72%, contra os 20-40% da co-infecção. Para MOIs elevadas, a diferença entre as duas estratégias não é significativa.

No Capítulo 7 discutem-se as implicações dos resultados apresentados nos outros capítulos. Resumem-se os principais objectivos alcançados

xxiii

durante a realização desta tese, identificam-se as maiores falhas nas ferramentas computacionais desenvolvidas nos capítulos anteriores para a otimização do sistema de baculovírus/células de insecto e definem-se estratégias que potenciem uma produção de RLP mais eficiente.

Em resumo, esta tese fornece uma serie de modelos computacionais e algoritmos que representam um grande passo na implementação de um sistema de baculovírus/células de insecto virtual. Através de simulações e códigos de optimização numéricos, novos e melhores esquemas de optimização para produção de partículas virais, proteínas virais e/ou estruturas proteicas complexas tais como as VLPs podem ser desenvolvidas.

## ***Table of Contents***

### **CHAPTER 1: Introduction**

<b>1. Introduction.....</b>	<b>3</b>
1.1. The baculovirus/insect cell system at a glance.....	4
1.2. Bioengineering issues and caveats .....	7
<b>2. Virus-like particles as vaccines: the proof of concept.....</b>	<b>12</b>
2.1. VLPs of structurally simple viruses .....	13
2.2. VLPs with lipid envelope.....	13
2.3. VLPs with multiple-protein layers.....	15
2.4. Chimeric VLPs.....	16
<b>3. Types of mathematical models .....</b>	<b>17</b>
3.1. Parametric models.....	17
3.2. Non-parametric models .....	19
3.3. Hybrid semi-parametric models .....	20
3.4. Outline of modelling approaches in insect cell cultures .....	21
<b>4. RLPs production: a Systems Biology perspective .....</b>	<b>22</b>
4.1. The concept of Rotavirus-like particles .....	22
4.2. The fate of an insect cell upon baculovirus infection .....	24
4.3. Comparison of selected models .....	31
<b>5. Scope of the thesis .....</b>	<b>39</b>

### **CHAPTER 2: Recombinant baculovirus titration**

<b>1. Introduction.....</b>	<b>54</b>
<b>2. Materials and methods .....</b>	<b>57</b>
2.1. Cell culture and baculovirus amplification.....	57
2.2. Baculovirus titration methods.....	58
2.3. Statistical analysis .....	65
<b>3. Results.....</b>	<b>66</b>
3.1. TCID <sub>50</sub> for baculovirus with and without GFP .....	66
3.2. MTT assay.....	69
3.3. Alamarblue assay .....	71
3.4. Cell size assay.....	72

3.5. Growth cessation assay.....	73
3.6. Real time Q-PCR .....	74
3.7. Plaque assay .....	74
3.8. Flow cytometry assay .....	75
3.9. Statistical analysis.....	76
<b>4. Discussion.....</b>	<b>78</b>
4.1. Critical issues in baculovirus titration .....	78
4.2. Error assessment in baculovirus titration methods .....	82
4.3. Estimated accuracy of titers.....	82
4.4. Methods cost, time and error analysis .....	83

## CHAPTER 3: Stochastic infection and protein expression in the AcMNPV/Sf-9 system

<b>1. Introduction.....</b>	<b>90</b>
<b>2. Proposed optimisation method.....</b>	<b>92</b>
2.1. Stochastic infection model .....	93
2.2. A structured model for intracellular protein expression .....	96
<b>3. Case study: production of vp2.....</b>	<b>98</b>
3.1. Model calibration.....	98
3.2. Stochastic simulation .....	100
3.3. Population dynamics.....	102
3.4. Intracellular dynamics .....	103
3.5. Productivity assessment .....	105
<b>4. Conclusions and future perspectives .....</b>	<b>107</b>

## CHAPTER 4: Thermodynamic-equilibrium based RLP assembly

<b>1. Introduction.....</b>	<b>114</b>
<b>2. Materials and methods.....</b>	<b>116</b>
2.1. RLPs production, purification and analysis .....	116
2.2. <i>In vitro</i> assembly and disassembly assays .....	117
<b>3. A model for rotavirus-like particle assembly.....</b>	<b>118</b>
3.1. Zlotnick's equilibrium framework .....	120



3.2. Application to the self-assembly of RLP .....	122
<b>4. Results and discussion .....</b>	<b>132</b>
4.1. Pseudo-critical protein concentration.....	132
4.2. The effect of the number of association contacts .....	133
4.3. Simulation of RLP assembly.....	134
4.4. <i>In vitro</i> RLPs assembly experiments .....	137
4.5. Temperature dependency.....	138
4.6. Interpretation of <i>in vitro</i> disassembly data .....	141
4.7. Design of optimal thermodynamic conditions for global RLP assembly.....	142
<b>5. Conclusion .....</b>	<b>145</b>

## CHAPTER 5: Model-based assessment of multiple infection strategies

<b>1. Introduction.....</b>	<b>152</b>
<b>2. Model formulation .....</b>	<b>154</b>
2.1. Process description .....	154
2.2. Modelling assumptions .....	155
2.3. Baculovirus adsorption .....	156
2.4. Infected and healthy cell population .....	158
2.5. Replication of vDNA .....	159
2.6. Synthesis of mRNA .....	161
2.7. Viral protein synthesis .....	161
2.8. Temporal control.....	162
2.9. RLPs assembly.....	164
<b>3. Materials and methods .....</b>	<b>164</b>
3.1. Cell culture and media .....	164
3.2. Baculoviruses design and infection strategies .....	165
3.3. Infection procedure .....	165
3.4. Nucleic acid quantification .....	166
3.5. Viral protein analysis .....	167
3.6. RLPs purification, quantification and characterization .....	167
3.7. Parameter estimation algorithm.....	168
<b>4. Results and discussion .....</b>	<b>169</b>
4.1. Model calibration and validation .....	169
4.2. Effect of viral infection on cell death rate .....	170

4.3. The dynamics of intracellular vDNA .....	171
4.4. Virus budding .....	173
4.5. mRNA synthesis .....	174
4.6. Protein synthesis .....	176
4.7. Analysis of protein dynamics .....	178
<b>5. Conclusions .....</b>	<b>179</b>

## CHAPTER 6: Optimization of RLP production using the stochastic model

<b>1. Introduction.....</b>	<b>188</b>
<b>2. The mathematical model.....</b>	<b>189</b>
<b>3. Sensitivity analysis of model parameters .....</b>	<b>191</b>
<b>4. Optimization of RLPs production .....</b>	<b>200</b>
4.1. Single-infection strategies .....	200
4.2. Co-infection strategies .....	206
4.3. What to choose: single-infection or co-infection?.....	213
<b>5. Conclusions and future perspectives .....</b>	<b>214</b>

## CHAPTER 7: Discussion and conclusions

<b>1. Discussion.....</b>	<b>221</b>
1.1. MOI controllability .....	221
1.2. Model-based optimisation of the MOI for single protein expression.....	222
1.3. Model-based optimisation of RLP production .....	224
<b>2. Conclusions .....</b>	<b>230</b>
<b>3. Future work .....</b>	<b>233</b>

## Abbreviations

Abbreviation	Full form
<b>AcMNPV</b>	<i>Autographa californica</i> multiple nucleopolyhedrovirus
<b>Ass<sub>eff</sub></b>	Assembly efficiency
<b>BTV</b>	Bluetongue virus
<b>CCI</b>	Cell concentration at infection
<b>DIP</b>	Defective interfering particles
<b>Dis<sub>eff</sub></b>	Disassembly efficiency
<b>DLP</b>	Double-layered particles composed by vp2 and vp6
<b>GFP</b>	Green fluorescence protein
<b>HIV</b>	Human immunodeficiency virus
<b>HPV</b>	Human papillomavirus
<b>MOI</b>	Multiplicity of infection (number of virus <i>per</i> cell)
<b>MTT</b>	Microculture tetrazolium
<b>pfu</b>	Plaque forming units
<b>polh</b>	Polyhedrin
<b>PTM</b>	Post-translational modifications
<b>rBac</b>	Recombinant baculovirus
<b>RLP</b>	Rotavirus-like particles
<b>SD</b>	Standard deviation
<b>Sf-9</b>	<i>Spodoptera frugiperda</i> cells
<b>SLP</b>	Single-layered vp2 particles
<b>TCID<sub>50</sub></b>	End-point dilution
<b>TOH</b>	Time of harvest
<b>vDNA</b>	Viral DNA
<b>VLPs</b>	Virus-like particles
<b>VP</b>	Viral protein

## List of Figures

<b>Figure 1.</b>	Number and percentage of proteins approved as biopharmaceuticals.....	4
<b>Figure 2.</b>	Reported expression levels of some proteins in <i>Sf</i> -9 cells and comparison to other expression systems. ....	8
<b>Figure 3.</b>	Illustration of a RLP. ....	22
<b>Figure 4.</b>	Genetic/molecular regulatory network of RLPs production in the baculovirus/insect cell system. ....	25
<b>Figure 5.</b>	Ishikawa diagram of the main bioengineering challenges of RLP production in the baculovirus/insect cell system. ....	40
<b>Figure 6.</b>	Electron micrograph of negatively stained RLPs, DLPs, single-layered vp2 particles and vp6 tube-like structures ....	41
<b>Figure 7.</b>	Overview of recombinant baculovirus titration methods and time required to perform each method. ....	57
<b>Figure 8.</b>	TCID <sub>50</sub> method for baculovirus with GFP: effect of cell concentration at the time of infection and incubation time on viral titers. ....	67
<b>Figure 9.</b>	Assessment of positive/negative wells using the TCID <sub>50</sub> method. ....	69
<b>Figure 10.</b>	MTT assay for baculovirus from stock A: effect of cell concentration at the time of infection on TCID <sub>50</sub> /ml values. ....	71
<b>Figure 11.</b>	Cell size assay for baculovirus from stock A to D: effect of incubation time on recombinant baculovirus titers. ....	73
<b>Figure 12.</b>	Estimated accuracy of baculovirus titers. ....	83
<b>Figure 13.</b>	Cost analysis of baculovirus titration methods.....	84
<b>Figure 14.</b>	Network of heterologous protein production in AcMNPV/ <i>Sf</i> -9 system.....	91
<b>Figure 15.</b>	Process kinetics of vDNA replication, mRNA synthesis and viral protein production for the genes coding for vp <sub>2</sub> and vp <sub>6</sub> .....	100
<b>Figure 16.</b>	Frequency histogram of total vp <sub>2</sub> concentration for N <sub>trials</sub> = 100 and MOI = 0.1 virus.cell <sup>-1</sup> . ....	102
<b>Figure 17.</b>	Kinetic profile of uninfected and infected cells population. ....	103
<b>Figure 18.</b>	Intracellular dna <sub>2</sub> , rna <sub>2</sub> and vp <sub>2</sub> profile for MOI = 0.01, 0.1, 1, 10 and 40 virus.cell <sup>-1</sup> . ....	104
<b>Figure 19.</b>	Effect of MOI on vp <sub>2</sub> production levels.....	105
<b>Figure 20.</b>	Gene size and MOI effect on final recombinant protein production. ....	107
<b>Figure 21.</b>	The quasi-equivalence theory of icosahedral virus assembly.....	119
<b>Figure 22.</b>	Diagram of the juxtaposition of vp6 trimers in relation to vp2 dimers. ....	125

<b>Figure 23.</b>	SLP assembly efficiency as function of the initial vp2 concentration for different Gibbs free energy values. ....	133
<b>Figure 24.</b>	The concentration of assembly intermediates for three initial protein concentration scenarios .....	136
<b>Figure 25.</b>	The effect of initial protein concentrations and Gibbs free energy of subunit association on RLP assembly.....	144
<b>Figure 26.</b>	Schematic representation of the RLPs production process for single- and co-infection strategies .....	154
<b>Figure 27.</b>	Model predictions over measured variables: infected cells concentration, intracellular vDNA and mRNA concentration, and total viral protein concentration. ....	170
<b>Figure 28.</b>	Dynamics of infected cell population .....	171
<b>Figure 29.</b>	Dynamics of intracellular vDNA2, vDNA6 and vDNA7. ....	172
<b>Figure 30.</b>	Intracellular dynamics of RNA2, RNA6 and RNA7. ....	174
<b>Figure 31.</b>	Total (intracellular plus extracellular) protein dynamics of vp2, vp6 and vp7. ....	176
<b>Figure 32.</b>	Intracellular vp2, vp6 and vp7 concentrations. ....	177
<b>Figure 33.</b>	Dynamics of proteins stoichiometric ratios. ....	178
<b>Figure 34.</b>	Sensitivity analysis of uninfected, infected and death cells concentration with respect to model parameter perturbations .....	193
<b>Figure 35.</b>	Sensitivity analysis of intracellular and total virus concentration with respect to model parameter perturbations.....	195
<b>Figure 36.</b>	Sensitivity analysis of intracellular vDNA, RNA and protein concentration with respect to model parameter perturbations.....	197
<b>Figure 37.</b>	Sensitivity analysis of total protein concentration with respect to model parameter perturbations .....	199
<b>Figure 38.</b>	The effect of MOI in single-infection strategies: evaluating the intracellular and total RLP production and the concentration of infected cells carrying the genes coding for vp2, vp6 and vp7.....	201
<b>Figure 39.</b>	The effect of promoter strength on total RLP concentration and assembly efficiency for single-infection strategies at MOI=0.01 virus.cell <sup>-1</sup> .....	204
<b>Figure 40.</b>	The effect of promoter strength on intracellular protein expression for single-infection strategies at MOI=0.01 virus.cell <sup>-1</sup> .....	205
<b>Figure 41.</b>	The effect of promoter strength on intracellular and total RLP production for single-infection strategies: maximum increase in RLP production that can be achieved through genetic redesign for a selective range of MOIs .....	206

<b>Figure 42.</b>	The effect of MOI on intracellular and total RLP production, assembly efficiency and concentration of infected cells carrying the genes coding for vp2, vp6 and vp7 for co-infection strategies .....	207
<b>Figure 43.</b>	Comparison of model outputs for traditional co-infection strategies and model-designed strategies based on oriented-manipulation of the MOI of individual baculoviruses.....	211
<b>Figure 44.</b>	Comparison of model outputs for co-infection strategies where the genes coding for the three viral proteins are under the control of promoters with equal strengths and co-infection strategies maximizing RLP production based on the manipulation of the strength of each promoter .....	212
<b>Figure 45.</b>	Single-infection vs co-infection: the effect of MOI on model .....	213

## List of Tables

<b>Table 1.</b>	Comparison of different expression systems.....	6
<b>Table 2.</b>	VLPs developed for vaccine .....	14
<b>Table 3.</b>	Index to named recombinant baculovirus gene/gene products. ....	27
<b>Table 4.</b>	Mathematical models describing the baculovirus infection process. ....	33
<b>Table 5.</b>	Mathematical equations for titer estimation. ....	60
<b>Table 6.</b>	Comparison of different methods for baculovirus titration. ....	70
<b>Table 7.</b>	Summary of the evaluated methods for baculovirus titration.....	79
<b>Table 8.</b>	Mathematical model parameters .....	99
<b>Table 9.</b>	Computation time and accuracy.....	101
<b>Table 10.</b>	Assembly intermediates and factors describing the formation of single-layered vp2 particles. ....	124
<b>Table 11a.</b>	The assembly of the first vp6 structural subunit on top of the RLP vp2 layer. ....	127
<b>Table 11b.</b>	The assembly of vp6 subunits on top of SLPs. ....	128
<b>Table 12.</b>	The assembly of vp7 subunits on top of DLPs. ....	131
<b>Table 13.</b>	<i>In vitro</i> assembly and disassembly of RLPs. ....	139
<b>Table 14.</b>	The burden caused by rotavirus disease.....	152
<b>Table 15.</b>	Experimental setup for protein expression. ....	158
<b>Table 16.</b>	Baculovirus infection and trafficking parameters. ....	169
<b>Table 17.</b>	Viral protein expression parameters.....	169
<b>Table 18.</b>	Mathematical model equations.....	190





# Chapter 1

**I**ntroduction

## CONTENTS

<b>1. Introduction</b>	<b>3</b>
1.1. The baculovirus/insect cell system at a glance	4
1.1.1. The baculovirus	4
1.1.2. The insect cell	5
1.1.3. The expression level	7
1.2. Bioengineering issues and caveats	7
1.2.1. The promoter	7
1.2.2. Baculovirus titration	9
1.2.3. The culture mode	9
1.2.4. Key process parameters	10
1.2.5. The metabolism of insect cells	11
1.2.6. Protein quantification methods	11
<b>2. Virus-like particles as vaccines: the proof of concept</b>	<b>12</b>
2.1. VLPs of structurally simple viruses	13
2.2. VLPs with lipid envelope	13
2.3. VLPs with multiple-protein layers	15
2.4. Chimeric VLPs	16
<b>3. Types of mathematical models</b>	<b>17</b>
3.1. Parametric models	17
3.2. Non-parametric models	19
3.3. Hybrid semi-parametric models	20
3.4. Outline of modelling approaches in insect cell cultures	21
<b>4. RLPs production: a Systems Biology perspective</b>	<b>22</b>
4.1. The concept of Rotavirus-like particles	22
4.2. The fate of an insect cell upon baculovirus infection	24
4.3. Comparison of selected models	31
4.3.1. Non-infected cell growth	31
4.3.2. Binding and infection	32
4.3.3. Infected cell population	35
4.3.4. vDNA and mRNA	36
4.3.5. Protein synthesis	37
4.3.6. The assembly of VLPs	38
<b>5. Scope of the thesis</b>	<b>39</b>
<b>6. References</b>	<b>43</b>

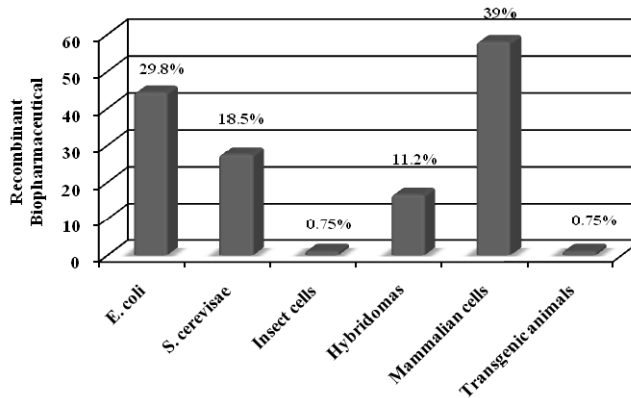
## 1. INTRODUCTION

The baculovirus/insect cell system is a two-stage process where insect cells are first grown to a desired cell concentration and then infected by genetically engineered baculoviruses for heterologous protein expression. This system is considered to be one of the most efficient eukaryotic gene expression systems. Over 500 recombinant proteins have been expressed to date, from viral, bacterial, fungi, mammalian and plant species, with yields ranging between 5 to 70% of the total intracellular protein content [1]; the concentrations can reach 500 mg.L<sup>-1</sup> [2]. Noteworthy is its capacity to express multiple proteins and to form multimeric protein complexes such as virus-like particles (VLPs) [3, 4] making this system a basic tool for protein-protein interaction research and VLP based vaccines development. Importantly, most of the proteins expressed in this system can undergo post-translational modifications (PTM) similar to the human glycosylation pattern [5, 6].

Despite its advantages, only two in 151 licensed protein-based biopharmaceuticals are produced using the baculovirus/insect cell system (Figure 1): the GlaxoSmithKline's human papillomavirus 16/18 cervical cancer vaccine Cervarix<sup>TM</sup> and the recombinant trivalent hemagglutinin vaccine against influenza virus, FluBlok<sup>®</sup>, from Protein Science Corporation.

The baculovirus/insect cell system can be used also for virus production. Applications of these viral particles go from high-throughput screening of gene functions [7] to drug delivery [8], *in vitro* assembly studies, design of antiviral drugs [9], bio-insecticides [10] or bio-weapons [11]. In material science and engineering, they are the building blocks for electronics, biosensors and chemistry [12]. Recently, baculoviruses have invited renewed interest by virtue of their potential use as a delivery system in

gene therapy [13, 14]. Although their replication is limited to arthropods, baculoviruses can enter in over 35 mammalian cell lines [15], delivering the vDNA into the nucleus. This is possible due to exogenous promoters, such as those derived from Rous sarcoma virus or cytomegalovirus.



**Figure 1.** Number and percentage of proteins approved as biopharmaceuticals

Evidence points out for promising applications of baculoviruses and protein-derived complexes to other areas of research in the near future such as nanotechnology where they can be used as tools for biomedical science [16].

## 1.1. The baculovirus/insect cell system at a glance

### 1.1.1. The baculovirus

Baculoviruses have a dsDNA genome of between 88 and 200 kbp with 90 to 180 genes [17, 18]. The genome is associated to a 6.5 kDa protein [19] and encapsidated into a rod-shaped nucleocapsid formed by a 39 kDa protein [20]. The length of the nucleocapsids varies from 200 to 400 nm, depending on the genome size, whereas the width ranges between 25 and 50 nm [21]. These nucleocapsids are packed within a lipoprotein envelope to form the virus particle.

The insertion of a foreign gene occurs normally at the polyhedrin (*polh*) or

*p10* gene locus since these are not essential genes for vDNA replication and are controlled by very strong promoters. The size of insertions can be as large as 100 kb, meaning that not only single but also multiple genes can be inserted and expressed simultaneously [22].

Baculoviruses are stable for 3-4 months when re-suspended in culture medium and stored at 4°C. For long term storage, -80°C or liquid nitrogen is recommended; additives such as glycerol, dimethyl sulfoxide and sucrose may avoid loss of infectivity. At -20°C, viral titers decrease significantly due to virus aggregation [23].

### **1.1.2. The insect cell**

Typical insect cell cultures do not require CO<sub>2</sub> for growth, can resist temperature fluctuations between 18 °C and 35 °C and osmolarities between 300 and 380 mOsm (a significant reduction in protein expression is observed if osmolarities then increase by 30mOsm) [24]. The optimum pH is around 6.2-6.3; it may reach 5.9 in infected cultures (Table 1).

Insect cells can be cultivated in static (T-flasks) or in suspension culture systems (erlenmeyer, shaker flasks, spinners or bioreactors) to high cell densities. They can be cultured indefinitely although for passages beyond the 50<sup>th</sup>, cells began to show morphological changes. Cells are easily adapted to suspension cultures, either in media containing serum or in serum free media, and are normally stored at -80°C (working cell bank) or in liquid nitrogen (long-term storage).

Insect cells, infected or not, are sensitive to shear stress generated while sparging or by bubble entrainment during agitation [25]. The addition of non-ionic copolymers (e.g. pluronic F-68) is essential as they lower the culture medium surface tension, impeding the attachment of cells to bubbles [26], and interact with the cell membrane, increasing its rigidity and

Table 1. Comparison of different expression systems

Characteristics	E. coli	Yeast	Insect cells	Mammalian cells
Cell doubling time (h)	0.5	1.5	18-24	24
Cell size (µm)	0.8-2	3-10	15-21	12-20
Dry weight (g/cell)	3×10 <sup>-13</sup>	10 <sup>-11</sup>	6×10 <sup>-10</sup>	3-6×10 <sup>-10</sup>
Process cell density (cell/ml)	2.3-3.3×10 <sup>11</sup>	> 1.9×10 <sup>10</sup>	1-14×10 <sup>6</sup> (A)	1-20×10 <sup>6</sup>
Cell density (g <sub>cell</sub> /L) (dry weight)	70-100	> 185	0.6-1.8	0.3-12
Product yield (mg <sub>prod</sub> /L)	50-1000	50-5000	0.5-500	0.1-100
Specific productivities (pg <sub>prod</sub> /cell)	0.0002-0.003	0.003-0.3	0.5-167	0.1-5
Process time	3-20 h	6-120 h	3-5 days	6-12 days
Productivity (mg <sub>prod</sub> /L/day)	400-1200	200-1000	0.2-100	0.02-8
Protein sizes (kDa)	< 50 (B)	> 50	> 50	> 50
Production cost (C)	Low	Low-moderate	Moderate	High
Scale-up	Easy	Moderate	Moderate	Complex
Purification	Easy	Moderate	Moderate	Complex
Post-translational modifications				
Proper protein folding	No (refolding required)	No (refolding required)	Yes	Yes
Proteolytic cleavage	Seldom occur	Seldom occur	Yes	Yes
Glycosylation	No	High Mannose	Simple (D)	Complex
O <sub>2</sub>	With or without	With or without (E)	Yes	Yes
CO <sub>2</sub>	With or without F	Not required	Not required	Required
Carbon source	Organic compounds or CO <sub>2</sub>	Glc, Fru, Mal, Suc	Glc (also Suc and Mal)	Mainly Glc and Gln
pH	6-7	Neutral (slightly acid)	6.2-6.3	7.2-7.4
Optimum temperature (°C)	37 (range 22-49°C)	30-37 (range 10-37°C)	27-29 °C	37
Recommended use	Protein antigens and standards, functional proteins, vaccines, bioremediation and immobilized enzymes	Model for genetics and cell biology, expression of glycosylated proteins, vaccines (alternative to insect cells)	Recombinant glycoproteins for many biochemical applications (vaccines, assay standards) and nanomedicine	Recombinant glycoproteins, tissue culture and engineering (functional and PTM studies, protein characterization) and vaccines

A – Batch mode (cell density can reach up to 57×10<sup>6</sup> cell.ml<sup>-1</sup> in fed-batch mode [27]); B – Normally, E.coli is incapable of efficiently producing very long or very short proteins, although the successful expression of a 210 kDa protein has been achieved [28]; C – Production cost includes costs associated with medium, bioreactors, purification process, possibility of contaminations (higher for eukaryotic systems than prokaryotic) and instrumentation to control and monitor the production process; D – does not contain penultimate galactose or terminal sialic acid; E – Anaerobic Yeast strains are very rare (facultative anaerobes); F – It is possible to control the growth of *Escherichia coli* by CO<sub>2</sub> [29]; Glc – Glucose; Fru – Fructose; Mal – Maltose; Suc – Sucrose; Gln – Glutamine; PTM – Post-translational modifications.

its resistance to hydrodynamic forces.

The control of the percentage of dissolved  $O_2$  ( $pO_2$ ) is essential to avoid oxygen limitation or excess, thereby inhibiting the synthesis of proteases or oxidative damage to proteins [30]. The specific growth rate of insect cells is not significantly affected by  $pO_2$  between 5% and 100%; in contrast, protein expression is constrained to  $pO_2$  values around 25-30% [31].

### **1.1.3. The expression level**

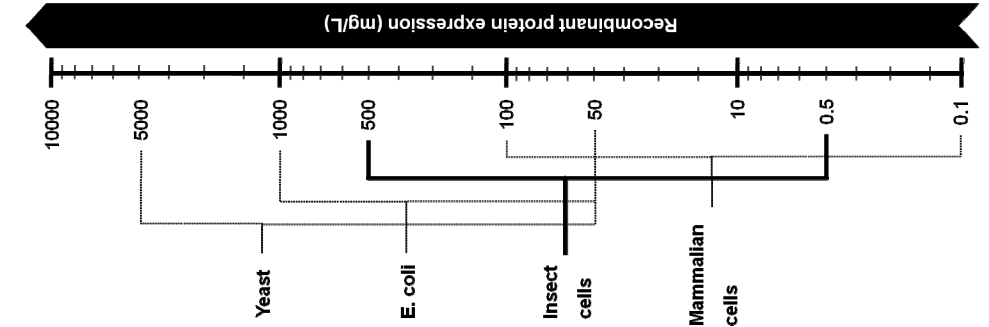
Recombinant protein yields may reach hundreds of milligrams *per* liter, depending on the culture mode (batch, fed-batch, continuous and perfusion) and the target protein (glycosylated or not) (Figure 2). These high expression levels are explained by the ability of the virus to shut off the transcription of early host genes and the consequent allocation of the cellular transcriptional and translational apparatus for the expression of the heterologous gene(s).

## **1.2. Bioengineering issues and caveats**

### **1.2.1. The promoter**

The combination of promoter's strength with the correct time for its activation is essential to high protein yields [32]. Very strong promoters commonly overwhelm the processing capacity of the endoplasmic reticulum, thus reducing final yields. In addition, with very late promoters, the expression of proteins occurs towards the end of the culture when proteases negatively impact on correct protein PTM. On the other hand, early promoters induce lower protein productivities as in most cases the enzymes and transcriptional factors necessary for protein expression are not yet active at such early stages.

Yield	Protein	Application	Ref.
> 500	Polyhedrin	Protection of baculoviruses from degradation	[33]
> 500	$\beta$ -galactosidase	Genetics, molecular biology and other life sciences	[34]
350	Human rotavirus vp6	Vaccine candidate against rotavirus disease	[35]
> 100	Chloramphenicol acetyltransferase	Reporter system to measure the level of a promoter expression	[36]
70	Tumor necrosis factor	Cell tumor regression activity	[37]
68	Hemagglutinin (HA1) from avian influenza H5N1	Vaccine for avian influenza H5N1	[38]
50	b-interferon	Multiple sclerosis	[39]
45	Granulocyte macrophage-colony stimulating factor	Infections related to bone marrow transplants	[40]
38	Tyrosine kinase ABL	Potential drug-target protein	[41]
35	Human parvovirus B19-VLP (virus-like particle)	Diagnostics and vaccination	[42]
25	Tyrosine kinase KDR	Potential drug-target protein	[41]
21.5	Viral protein VP2 of Infectious Bursal Disease Virus (IBDV)	Stimulation of immune response against IBDV	[43]
21-22	Interleukin	Stimulates the normal function of the immune system	[44]
20	Recombinant human vascular endothelial growth factor	Normal and pathologic angiogenesis	[45]
20	Nerve growth factor	Neurological diseases	[46]
10	Human Cyclin-dependent Kinase 2	Cell cycle regulation	[47]
10	Erythropoietin	Therapeutic agent	[48]
10	Epstein-Barr virus gp350/220	Cytotoxicity and antiviral immunosurveillance	[49]
5-10	West Nile virus envelope protein (WNV-E)	Candidate human and animal vaccine antigen	[50]
5	(His)6-ABL kinase protein	Chronic myeloid leukemia	[51]
5	Human LKB1	Tumor-suppression protein	[52]
4-10	Recombinant human amelogenin protein	Regeneration of periodontal tissues	[53]
4	Rotavirus-like particles of vp2/vp6/vp7	Vaccine candidate against rotavirus disease	[54]
3-5	Merozoite surface protein 1	Most promising malaria vaccine candidate	[55]
3	Recombinant human (rh)-renin	Regulates blood pressure in arterial smooth muscle cells	[56]
2-3	Tissue plasminogen activator	Acute myocardial infarction	[44]
2.5	Interleukin-5	Therapy of autoimmune disease and cancer	[57]
2.4	Glucocerebrosidase	Gaucher's disease	[58]
2.3	Porcine Parvovirus-VLPs	Vaccination against porcine parvovirus	[59]
2-20.5	Complement receptor	Pathogens detection	[60]
< 2	Serine/threonine kinase PKCh	Potential drug-target proteins	[41]
1	Epidermal growth factor receptor	Clinical cancer-related applications	[61]
1	Granulocyte-colony stimulating factor	Chemotherapy induced neutropenia	[62]
0.5-3.2	Rotavirus double layered particles of vp2/vp6	Vaccine candidate against rotavirus disease	[63]
0.5-2	Tissue factor	Coagulation and cytokine	[64]



**Figure 2.** Reported expression levels of some proteins in *Sf-9* cells and comparison to other expression systems.



Although previous reports show that weaker and earlier promoters induce correct PTM [65], early promoter-based vectors with productivities similar to *polh* and *p10* promoters are difficult to find.

### **1.2.2. Baculovirus titration**

To optimize conditions for efficient recombinant protein expression, it is important to measure as accurately as possible the titers of baculovirus stocks. Traditionally, plaque assays have been used for titer determination. Other methods such as endpoint dilution, microculture tetrazolium assay, growth cessation and cell size assay, alamarblue assay or electron microscopy have been proposed to aid the determination of virus titers. However, most of these methods are complicated to handle, time-consuming and labor intensive [66]. To shorten the titration times, techniques based on flow cytometry [67] or real-time polymerase chain reaction [68] were developed, showing reliable and comparable results to those determined by traditional titration assays. Despite these improvements, baculovirus titration remains a challenging task as existing methods are unable to differentiate between infectious and non-infectious particles, thus rendering difficult the estimation of accurate titers.

### **1.2.3. The culture mode**

Proteins can be expressed in batch, fed-batch, continuous or perfusion mode. In batch cultures, the control of toxic products formation and nutrients consumption is essential to guarantee the normal cellular growth and protein expression [69]. To avoid process limitations, selective addition of nutrients and/or amino acids may be performed using fed-batch strategies [70]. The yields are significantly improved but the scale-up becomes expensive since culture medium is inefficiently used. The continuous system presents a short throughput time and a small number of

production steps. However, continuous reactors operated for long periods generate high amounts of defective interfering particles (infectious viruses lacking viral genome) [71]; also, the complexity of the bio-product produced is generally low. Perfusion mode enhance cellular growth and protein expression [72], but require large volumes of media and cell-medium separation devices that impact negatively on cellular growth rate, protein expression and production cost.

#### **1.2.4. Key process parameters**

Process parameters such as the multiplicity of infection (MOI - number of virus *per* cell), time of harvest (TOH) and cell concentration at infection (CCI) strongly influence protein expression.

The optimal MOI is dependent on the target product (single or multiple protein expression), the production strategy (single- or co-infection) and the viral stock. For single protein expression, low MOIs normally induce higher volumetric protein productivities than high MOIs as a consequence of higher infected cell concentrations [73]. In theory, this difference could be balanced with the increase in CCI. However, cell specific productivity decreases as CCI increases due to the “cell density effect” [69]. For multiple protein expression, the self-assembly of protein complexes is highly correlated with the production strategy; thus, several combinations of MOI and CCI may maximize product formation.

Optimal TOH is normally between 72-120 hpi (40 and 70% of cell viability) [31]; outside these boundaries, protein degradation is increased and downstream processing becomes more difficult as contaminant proteins (host and viral) are released to the extracellular medium along with the product.

### **1.2.5. The metabolism of insect cells**

The preferred carbon source of insect cells is glucose; once exhausted, cell metabolism changes and glutamine becomes an important energy source [69]. Amino acids fulfil the nitrogen requirements of insect cells [74, 75]. Deprivation of lipids causes cell degeneration and the formation of non-infectious viruses [76]. Cholesterol is essential for membranes synthesis [77]. Vitamins are important for maintenance of insect cells since they act as cofactors for key metabolic enzymes [77].

Accumulation of toxic products is not observed at adequate nutrient and, most importantly, oxygen supply [75, 78]. At glucose and glutamine limitation levels, accumulation of ammonia is observed, followed by a significant reduction in cell growth rate [79]. Lactate is generated under oxygen deprivation but does not affect cellular growth until concentrations higher than 15 mM; lactate is consumed upon carbohydrates exhaustion causing the pH to increase towards the end of the culture [80].

Nutritional requirements of insect cells upon baculovirus infection have been generally considered to remain unchanged; cellular energetics seems to be the main limitation to baculovirus and/or protein production [69].

### **1.2.6. Protein quantification methods**

The quantification of protein complexes and constitutive viral proteins is traditionally performed by transmission electron microscopy, SDS-PAGE, Western blot, ELISA and BCA protein quantification assays [68]. These semi-quantitative methods do not differentiate between unassembled and assembled proteins; also, they are time-consuming, with low sensitivity and require significant amounts of sample. As a consequence, protein yields are frequently over or under estimated. Recently developed methods such as gel permeation HPLC [63], sodium dodecyl sulfate-capillary gel

electrophoresis [81], MALDI-TOF MS [82] and capillary zone electrophoresis [83] may allow a simple, fast and high performance quantification of unassembled and assembled proteins in purified and bulk samples.

## **2. VIRUS-LIKE PARTICLES AS VACCINES: THE PROOF OF CONCEPT**

Virus-like particles offer several advantages as antigens and immunogens. First, compared to individual proteins or peptides, VLPs present conformational epitopes more similar to the native virus; hence, antibody reactivity or immune system response is expected to be significantly improved. Secondly, because VLPs are non-infectious (they lack the viral genome), attenuation or inactivation is not required; this is particularly important as epitopes are commonly modified by inactivation treatments [84]. Thirdly, using molecular biology methods it is possible to adapt one or more antigens to these multimeric protein structures for broader and more efficient protection. Finally, VLPs resembling the organizational structure of native viruses imply that lower doses of antigen are required to elicit a protective response when compared to subunit vaccines.

The most popular expression system for the production of VLPs is the yeast system (*Saccharomyces cerevisiae* or *Pichia pastoris*) due to its easy protein expression, ability to scale-up and cost of production. However, appropriate PTM such as protein glycosylation and correct protein folding, protein assembly and codon optimisation may dictate alternative production systems. Within those, the baculovirus/insect cell system, mammalian cell lines (CHO, BHK-21, Vero E6, HepG2 or HeLa cells) either transiently or stably transfected or transduced with viral expression vectors, bacteria (*E. coli*) and various species of plant cells (*Solanum tuberosum* – potato,

*Lycopersicon esculentum* – tomato, *Nicotiana benthamiana* – tobacco), have been receiving special attention.

To date, different types of VLPs have been developed for vaccine proposes: 1. VLPs of structurally simple viruses; 2. VLPs with lipid envelope; 3. VLPs with multiple-protein layers; 4. Chimeric VLPs (Table 2).

### **2.1. VLPs of structurally simple viruses**

VLPs of structurally simple viruses are normally composed by one protein. One example is the human papillomavirus (HPV)-VLP. Although the native virus contains the major and minor capsid proteins of HPV, L1 and L2 respectively, the HPV-VLP is formed just by the L1 protein organized in 72 pentameric capsomers. Porcine parvovirus (PPV)-VLP is another example; PPV-VLPs are formed by a single protein, VP2, the major structural protein.

Norwalk virus (Nv)-VLPs expressed in insect cells are used as a source of diagnostic antigen to monitor disease outbreaks since the native virus production has limited growth in cell culture. Nv-VLPs have shown to be effective at stimulating IgG, IgA and humoral responses in mice [85]. Preliminary phase I trials in humans have confirmed that they are safe and effectively stimulate IgG and IgA responses [86].

### **2.2. VLPs with lipid envelope**

Many pathogenic viruses are surrounded by an envelope consisting of host cell membrane lipids and proteins, and viral glycoproteins. These proteins are the targets of neutralizing antibodies and essential components of a vaccine. Due to the inherent properties of the lipid envelope, assembly of enveloped VLPs is technically complex.

Table 2. VLPs developed for vaccines

Virus	Expression system	Recombinant proteins	Size	Product status	References
AAV	B/C	VP1, VP2, VP3	20-25 nm	Development	[87]
BTV	B/C	VP2, VP3, VP5, VP7	69 nm	Preclinical	[22]
Ebola virus	MC <sup>(A8)</sup> and B/C	VP40 and glycoprotein	65-70 nm	Preclinical	[88]
Enterovirus 71	B/C	VP0, VP1, VP2	25-27 nm	Preclinical	[89]
H1N1	B/C	HA, NA, M1	80-120 nm	Phase IIa (Novavax, Inc.)	[90]
H1N1	Transgenic plants	HA	80-120 nm	Preclinical	[91]
H3N2	B/C	HA, NA, M1	80-120 nm	Phase IIa (Novavax, Inc.)	[92]
H5N1	B/C	HA, NA, M1	80-120 nm	Phase I/IIa (Novavax, Inc.)	[93]
H5N1	Transgenic plants	HA	80-120 nm	Phase I (Medicago, Inc.)	[94]
HBV	<i>S. cerevisiae</i>	HBsAg	22 nm	Licensed (Recombivax HB <sup>®</sup> , Merck & Co. and Engerix-B <sup>®</sup> , GlaxoSmithKline)	[94, 95]
HBV	Yeast	HBsAg	18 nm	Licensed (Hepavax-Gen <sup>®</sup> , Berna Biotech Korea Corp)	[96]
HBV	MC	PreS1, preS2 and HBsAg	22 nm	Licensed (Bio-Hep-B <sup>®</sup> , BioTech Gen Corp)	[97]
HCV	B/C	Core, E1 and E2 proteins	40-60 nm	Preclinical	[98, 99]
HIV	MC <sup>(A8)</sup> and B/C	Pr55gag and envelope proteins	100-120 nm	Preclinical	[100]
HIV	MC-VVES	Gag and pol proteins	100-150 nm	Development	[101]
HIV	<i>S. cerevisiae</i>	Gag protein	108-138 nm	Development	[102]
HPV	<i>S. cerevisiae</i> and B/C	L1 capsid protein	30-50 nm	Licensed (Gardasil <sup>®</sup> , Merck & Co. and Cervarix <sup>™</sup> , GlaxoSmithKline)	[103]
HPV	<i>E. coli</i>	L1 capsid protein	50-60 nm	Preclinical	[104]
HPV	MC-VVES	L1 capsid protein	40-50 nm	Development	[105]
H-PV	B/C	VP1, VP2	22 nm	Phase I/II (NIH / Meridian life Science)	[106]
IBDV	MC-VVES and B/C	VP2, VP3, VP4	60-65 nm	Development	[107]
Marburg virus	MC <sup>(A)</sup> and B/C	VP40 and glycoprotein	50-100 nm	Preclinical	[108]
Nv	B/C	Capsid protein	38 nm	Phase I (LigoCyte Pharmaceuticals, Inc.)	[86]
Nv	Transgenic plants	Capsid protein	23-38 nm	Preclinical	[109]
Poliovirus	B/C	VP0, VP1, VP3	27 nm	Development	[110]
PPV	B/C	VP2	22 nm	Preclinical	[111]
Rotavirus	B/C	VP2, VP6, VP7	70-80 nm	Preclinical	[68]
RSV	<i>E. coli</i>	Gag protein	70 nm	Development	[112]
SIV	B/C	Pr56gag, envelope protein	130 nm	Preclinical	[113]

<sup>A</sup> Transient transfection; <sup>B</sup> Stable cell line; **AAV** – adeno-associated virus; **B/C** – baculovirus/insect cells; **BTV** – bluetongue virus; **HA** –hemagglutinin; **HBV** – hepatitis B virus; **HBsAg** – Hepatitis B surface antigen; **HCV** – hepatitis C virus; **HIV** – human immunodeficiency virus; **HPV** – human papillomavirus; **H-PV** – human papillomavirus; **IBDV** – infectious bursal disease virus; **M1** – matrix 1; **NA** – neuraminidase; **MC** – mammalian cells; **Nv** – norwalk virus; **PPV** – porcine parvovirus; **RSV** – rous sarcoma virus; **SIV** – simian immunodeficiency virus; **SV** – vaccinia vector expression system.

The first VLP vaccine to be produced consisted of spherical particles with diameters between 17 nm and 25 nm formed by the surface antigen of the hepatitis B virus (HBsAg) co-assembled with host cellular membranes. Following that, other enveloped VLPs have been successfully developed. The hepatitis C virus-VLP and VLPs of Ebola and Marburg virus are some examples. The first has been tested in mice and baboons and shown to be effective at stimulating both cellular and humoral immune responses [98, 99]. The latter ones have shown to protect small laboratory animals as well as non-human primates against lethal challenge by Ebola or Marburg viruses [114].

Simian immunodeficiency virus (SIV)-VLPs, human immunodeficiency virus (HIV)-VLPs, Rous sarcoma virus-VLPs, and VLPs for Influenza A H1N1, H3N2 and H5N1 viruses are also part of the group of successfully produced VLPs with lipid envelope. Although none of the retrovirus derived VLPs have gone into clinical trials yet, initial experiments with HIV-VLP [115] and SIV-VLP [116] in animal models look promising.

### **2.3. VLPs with multiple-protein layers**

VLPs composed by multiple proteins are more challenging to produce than those formed by only one protein due to the site of protein expression. Proteins encoded by multiple discrete mRNAs tend to be differently localized in the cell, which may significantly affect the assembly process. Thus, it is essential to guarantee that all proteins are expressed in the vicinity of each other and within the same cell.

These complex VLPs have been efficiently produced for various members of the *Picornaviridae* (poliovirus and enterovirus 71), *Birnaviridae* (infectious bursal disease virus) and *Reoviridae* families (bluetongue virus

and rotavirus); included in this VLP's category are also the adeno-associated virus type-2 VLPs.

Intact and biologically active bluetongue virus (BTV)-VLPs can be produced in insect cells by co-infection using two bicistronic baculovirus, one encoding VP3 and VP5 and the other encoding VP2 and VP7, or by single-infection using multicistronic baculoviruses that express simultaneously all four structural proteins of the BTV (VP2, VP3, VP5 and VP7). With the use of four-gene baculovirus vectors, the possibility of some of the cells not being co-infected by the two bicistronic baculovirus is overcome. For this reason, single-infection strategies induce the formation of more BTV-VLPs than co-infection strategies.

Biologically functional rotavirus-like particles (RLPs) are normally expressed in the baculovirus/ insect cell system using co- or single-infection strategies. In co-infection, cells are infected with three monocistronic baculoviruses, each one coding for vp2, vp6 or vp7. In single-infection, cells are infected with tricistronic baculoviruses coding for the three viral proteins simultaneously (vp2, vp6 and vp7). Similar to BTV-VLPs, single-infection strategies appear to be advantageous over co-infection [68]. In animal models, RLPs have shown to induce a robust antibody response and protection when formulated with a potent adjuvant and administered intramuscularly [3, 117]. Even intrarectal immunization, inducing a local mucosal response, has been reported as sufficient for protection from rotavirus infection in animal trials [118].

## **2.4. Chimeric VLPs**

A forthcoming application of VLPs is in the generation of immune responses against foreign protein epitopes by fusing or by coupling them to VLPs of different origins, resulting in a chimeric VLP. These VLPs are able



to carry multiple foreign epitopes [119] and, with the incorporation of tags (e.g. polyhistidine), allow easy single-step product recovery [120]. The vaccine against HPV infection is an example of a chimeric VLP, in which the L2 protein epitopes are inserted into the L1 protein to confer protection against a broader range of HPV types [121].

### 3. TYPES OF MATHEMATICAL MODELS

In biological sciences, mathematical models constitute a compact means of describing complex mechanisms and a tool to predict the behavior of a system to variations in previously unstudied parameters. Mathematical models can be divided in three categories: 1. parametric models, which are based on *a priori* knowledge; 2. non-parametric models, whose structure is determined from data; 3. hybrid semi-parametric models, which combine parametric and non-parametric components.

#### 3.1. Parametric models

A parametric or white-box model is based on First Principles and mechanistic knowledge. A functional relationship is derived supported by the *a priori* knowledge relating the different variables. The main bottleneck of parametric models is sometimes the low predictive capacity due to the over-simplifications of the underlying biological system.

White-box models can be classified on the basis of their use (descriptive/optimization), time dependency (static/dynamic), number of time-related data (discrete/continuous), degree of randomness (deterministic/stochastic), intracellular processes (structured/non-structured) and the detail in the description of the cell population (segregated/non-segregated).

Descriptive models are used to merely assess mathematical relationships; an example is the study of the effect of temperature and pH on microbial growth [122]. In contrast, optimization models are used to find optimal solutions; a successful application is the production optimization of a recombinant fusion glycoprotein in Baby Hamster Kidney cells [123].

Static models describe the state of the components of a system, not accounting for the element of time [124]. In contrast, dynamic models take into account the time and are typically represented by ordinary or partial differential equations [43, 125]. Within dynamic models, there are discrete and continuous models; the difference between both is based on the number of time-related data. If the number is finite, the discrete model applies quite naturally, no matter whether the data belong to time-points or represent changes over time-intervals. Discrete modeling approaches have been applied to protein expression dynamics [126], gene regulatory networks [127] and epidemiology [128]. For near-infinite data, the continuous model fits naturally.

Deterministic models are those for which the relationships between variables are known with certainty [129]; hence, a given set of initial conditions will always generate the same output. Models where those relationships are not known are said to be stochastic models [73]. A stochastic process uses a range of values for variables in the form of probability distributions; randomness is present, even when given an identical set of initial conditions.

Unstructured models completely disregard the intracellular state. They only consider extracellular state variables; the reaction rates are described empirically as functions of extracellular conditions [130]. Structured models include details about the intracellular pools, namely concentrations of metabolites, enzymes, DNA, mRNA, amongst others [43, 73].

Unsegregated models assume that a culture is composed by a collection of cells which have the same characteristics and behave equally [131]. On the contrary, segregated models treat each cell as independent, and a population as a collection of such distinct cells. Such cell heterogeneity is normally modelled using probability functions such as the Poisson or Weibull distribution [43, 132]. These models are useful to describe cells at different phases of the cell cycle and cellular ageing, different morphological types of cells and interactions between different cells.

### 3.2. Non-parametric models

With the advent of high-throughput methods, it became clear that mechanistic modelling formalisms were unable to cope with the huge amounts of data of large scale molecular networks. Thus, non-parametric models were developed and used to analyze such complex networks of interactions.

A non-parametric or black-box model does not use *a priori* knowledge about the system at hand. It is necessary to estimate not only the values of model parameters but also to identify the functional relationships between variables and parameters. When *a priori* information is available, a set of functions can be used to reduce the “dimension” of the model. Since that is not the case, universal approximation functions with undermined number of parameters must be hypothesized in order to cover all the main biological mechanisms. This increases the number of empirical parameters and also the number of measurements required for their estimation.

Non-parametric models are able to deal with complex relationships between inputs and outputs in systems where little is known about the biochemical phenomena occurring within the cell, or to find patterns in data. However, this wider applicability has a cost in the robustness of the model;

in cases where a parametric model is more appropriated, non-parametric models have less predictive capacity, i.e. a larger sample size may be required to draw conclusions with the same degree of confidence.

An often used black-box model for bioprocess development is the multilayered artificial neural network (ANN) [133]. Basically, it consists of an interconnected group of artificial neurons, called nodes, organized into different layers and through which correlations between independent and dependent variables are established. The input layer receives the independent variables of the system while the output layer nodes generate the dependent variables. The inner layers are called hidden layers. Parameters such as the number of nodes in hidden layers or the number of hidden layers determine the structure of the network; this structure may vary as result of an adaptation to external or internal information that flows through the network during the “model’s learning phase”.

### **3.3. Hybrid semi-parametric models**

Hybrid semi-parametric models combine parametric models, based on First Principles and mechanistic knowledge, with empirical non-parametric models. They can be classified as parallel or serial hybrid structures [134, 135]. In the former case, the non-parametric model is combined in parallel with the full and insufficiently accurate mechanistic model, having access to the same input variables and adjusting the mechanistic model outputs. In the latter, the non-parametric model collects information from data and feeds it into the mechanistic parts of the parametric model that are not known, thereby filling the gaps. This approach is particularly appropriate to systems where large data sets are available and no direct mechanistic interpretation exists.

### 3.4. Outline of modelling approaches in insect cell cultures

Most of the present models for insect cell culture are parametric/unstructured models, describing only the growth, death and basic metabolic features of insect cells; the application of these models in model-based process control and process optimisation is therefore limited. To close the gap between extra- and intra-cellular compartments, structured models have been developed. These models describe the interactions between the different cellular pools as well as the extracellular environment using linear, algebraic and probability equations, ordinary or partial differential equations, traditional Michaelis-Menten equations, amongst many others. An illustration of the level of complexity these interaction networks can reach is the structured model proposed by Jang and co-workers (2000) [136]. In their work, a wide range of energy production and biosynthesis pathways were modelled; the concentration of 40 key metabolites in three regions (growth medium, cellular cytoplasm and cellular mitochondria) was considered. The model includes 1710 differential and algebraic equations and 390 model parameters. Due to the model's complexity and the number of parameters, automatic parameters' estimation was impossible; hence, selected heuristical parameters were tuned manually. This illustrates that structured models, although attractive due to their wide applicability, suffer from the shortcomings of being highly complex, computationally intensive and involving a large number of indeterminable parameters.

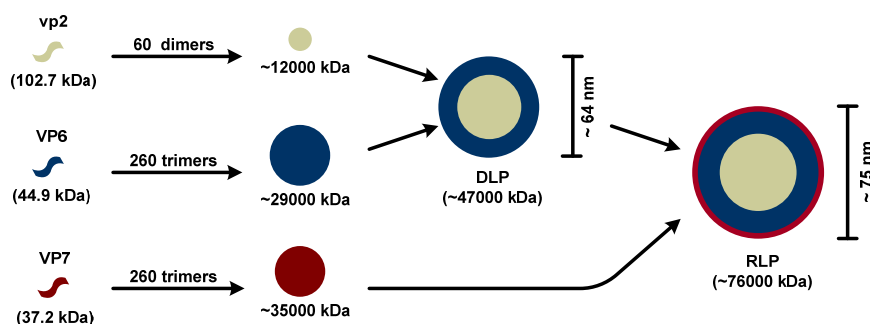
Recently, non-parametric and hybrid semi-parametric models have been emerging as powerful tools for the mathematical description of complex animal cell processes. However, none of these methods have been, to date, applied to the baculovirus/insect cell system for process control and optimization.

## 4. RLPS PRODUCTION: A SYSTEMS BIOLOGY PERSPECTIVE

Spurred on by the increase in the “omics” data and the advances in high-throughput experiments and bioinformatics, Systems Biology has emerged as a powerful tool to assist process engineers in accelerating the timeline for moving candidate biopharmaceutical products into flexible production platforms. In this work, a bottom-up Systems Biology approach was adopted for the optimization of RLP production, a vaccine candidate against rotavirus disease.

### 4.1. The concept of Rotavirus-like particles

RLPs are multi-protein complexes that mimic the organization and conformation of native non-enveloped rotaviruses but lack the viral genome. These spherical shaped particles are composed by three viral proteins (vp) of rotavirus arranged in a triple layered structure: a core of 60 dimers of vp2 – 102.7 kDa [137]; a middle shell of 260 trimers of the polymorphic protein vp6 – 44.9 kDa [138] and an outer layer of 260 trimers of a 37.2 kDa glycoprotein – vp7 [138] (Figure 3).



**Figure 3.** Illustration of a RLP: molecular weights and sizes of vp2, vp6, vp7, double-layered particles of vp2 and vp6 (DLP) and RLP are presented in kDa and nm respectively.

RLPs are vaccine candidates against rotavirus disease, a disease that, according to UNICEF/WHO report of October 2009 focused on diarrheal

disease (the second leading killer of children under 5), is responsible for more than 600,000 deaths in infants and children worldwide annually (source: <http://www.unicef.org>).

The main advantage of this vaccine candidate is the absence of viral genetic material; adverse effects such as intussusception, has happened in 1998 with the RotaShield<sup>®</sup> (Wyeth-Ayerst Laboratories) vaccine, reversion to virulence or mutations, are impossible to occur. In addition, it is possible to adapt one or more antigens to RLPs for broader and more efficient protection against a specific rotavirus serotype [139]. The fact that RLPs' resemble the structure of the native virus means that lower doses of antigens are required to elicit a protective response when compared to subunit vaccines. For these reasons, RLPs may constitute a viable alternative to commercial live attenuated vaccines (RotaTeq<sup>®</sup>, Merck & Co. and Rotarix<sup>®</sup>, GlaxoSmithKline), which have recently raised efficacy concerns in developing countries such as Bangladesh and South Africa [140].

RLPs are normally expressed in the baculovirus/insect cell system either using co-infection strategies, where cells are infected with three monocistronic baculoviruses enclosing the genes coding for vp2, vp6 or vp7, or single-infection strategies, where cells are infected with one tricistronic recombinant baculovirus enclosing the genes coding for the three viral proteins simultaneously. In single-infection strategies, different promoters may induce different expression levels at more convenient times for correct particle formation; hence, further optimizations are possible through vectors redesign at the promoters' level. On the other hand, in co-infection strategies, it is likely that productivities may be highly improved through manipulation of key process parameters (MOI, TOI and CCI); this poses an optimal control problem requiring an accurate dynamical model of the critical events involved in protein expression and RLP assembly.

## **4.2. The fate of an insect cell upon baculovirus infection**

RLP production in the baculovirus/insect cell system involves a number of intra- and extra-cellular steps, molecular interactions and genetic regulation mechanisms that determine the efficacy of the overall process (Figure 4).

The process can be divided into the following phases:

### **Phase 1. Virus infection and trafficking**

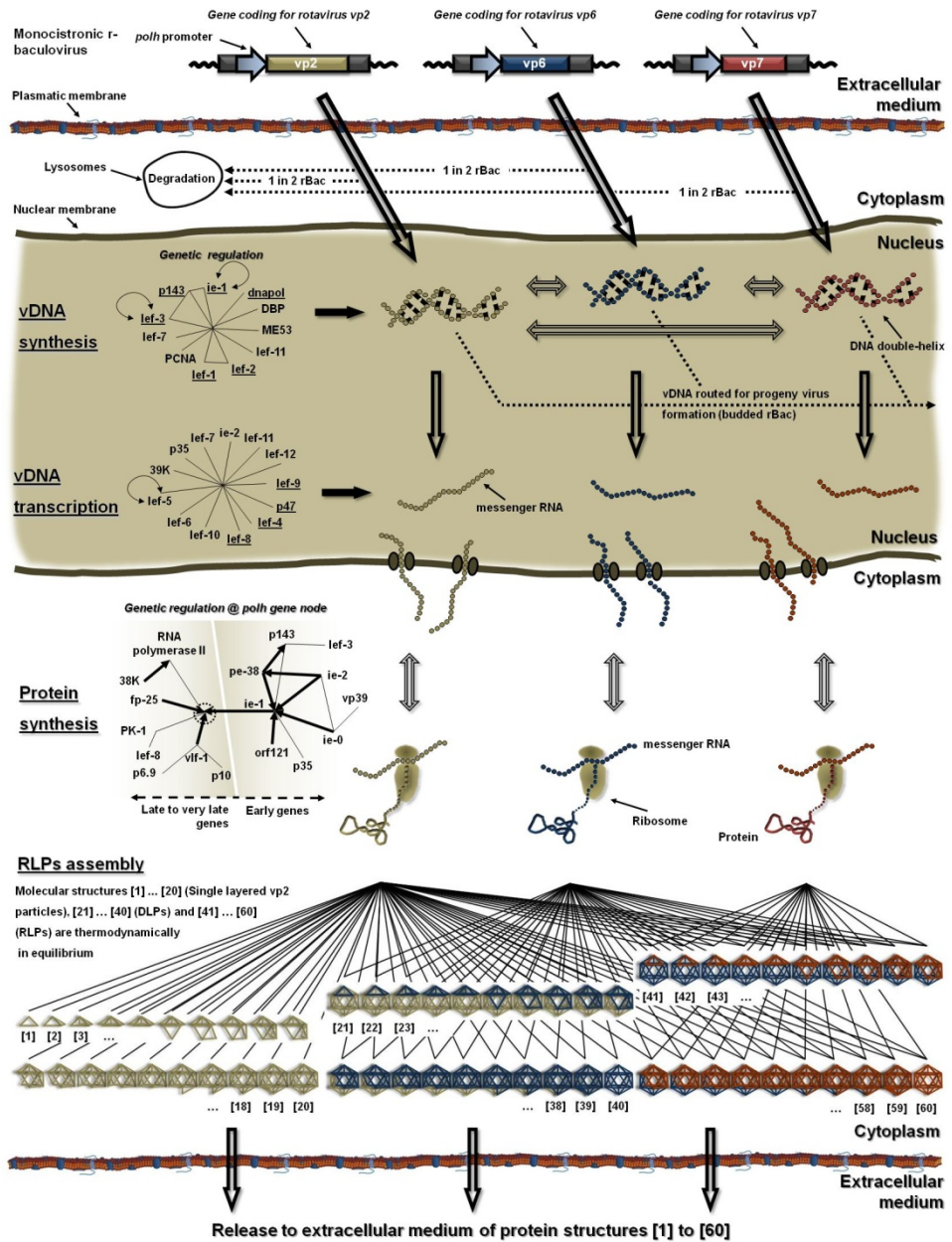
The process begins with the binding of monocistronic or tricistronic baculoviruses, enclosing the genes coding for rotavirus vp2, vp6 and vp7, to *Sf-9* cells membrane, a process mediated by cell surface receptors. About 50% of the viruses reach the host cell nucleus; the remaining viruses are degraded in lysosomes [141].

Once inside the nucleus, viruses and host cell compete for the control of the cell's well-organized and programmed transcriptional machinery, strictly regulated for carrying out host processes. To gain control, viruses selectively inhibit host cell transcription or eliminate host mRNA. Although encoding their own RNA polymerase and transcriptional factors, activators and enhancers, baculoviruses require the host RNA polymerase for the transcription of early genes, essential for initiation of vDNA replication, and expression of anti-apoptotic genes.

#### **Phase 1 - Stage 1. Early genes (immediate and delayed)**

The expression of early genes occurs immediately upon infection ( $\leq 2$  hpi), without requiring *de novo* synthesis of other virus gene products. The immediate early (ie)-1 transcriptional activator starts the transcriptional cascade, initiating the replication of baculoviruses.





**Figure 4.** Genetic/molecular regulatory network of RLPs production in the baculovirus/insect cell system. The arrows  $\Rightarrow$  and  $\Leftarrow$  represent transport of material and interaction between structures, respectively. In each genetic regulation networks, the line  $\text{—}$  and the arrows  $\longrightarrow$  and  $\longleftarrow$  indicate an interaction between two genes, the stimulation of one gene by a second gene and the stimulation of one gene by itself, respectively. The genes directly involved in vDNA replication and transcription are underlined.

Other genes involved in this process are the ie-0, ie-2, p143, 39K, anti-apoptotic p35 genes, viral DNA polymerase (dnapol) and pe-38 (Table 3).

### **Phase 1 - Stage 2. Anti-apoptotic genes**

It is upon the expression of ie-1 that increased programmed cell death (apoptosis) is observed. Two explanations are possible: 1. apoptosis is triggered after vDNA replication due to the expression of ie-1 and also late genes; 2. vDNA replication triggers apoptosis by damaging the host cell DNA or by disturbing the insect cell cycle. Apoptosis can be delayed by the expression of anti-apoptotic genes such as the p35 and the inhibitor apoptosis genes.

### **Phase 2. vDNA replication**

Replication of vDNA starts in the late phase (~ 6 hpi) after the expression of essential genes in the early phase; it relies on six viral genes: ie-1, p143, late expression factor (lef)-1, lef-2, lef-3 and dnapol. The dnapol is the enzyme involved in initiating vDNA replication; in its absence no vDNA replication is observed. The DNA binding proteins, lef-11, ME53, lef-7 and pcna, although not essential, are also involved in the process of vDNA replication.

### **Phase 3. Late gene expression**

The transcription of vDNA into mRNA and the expression of late genes occurs in this phase, concurrently with the onset of vDNA replication. There are 19 genes implicated or affecting late gene transcription [142]: four genes are components of the viral-encoded late RNA polymerase complex (lef-4, lef-8, lef-9 and p47); six genes are involved in vDNA replication (ie-1, lef-3, p143, lef-1, lef-2 and dnapol) and nine genes indirectly stimulate late transcription (p35, ie-2, 39K, lef-5, lef-6, lef-7, lef-10, lef-11 and lef-12).

**Table 3.** Index to named recombinant baculovirus gene/gene products.

Genes	Regulation	Function	References
ie-1 (E/LV/L) <sup>A</sup>	(+) p35, p143, 39K, ie-2, gp64, pcna and polh (s) dnapol, pe-38, ie-0, orf121 and hrs	vDNA replication; associated with the virogenic stroma; shut-off mechanism during the late phase of infection; thought to induce apoptosis; acts as the origin binding protein serving as a marker for enucleating the assembly of a replication complex	[143]
p143 (E) <sup>B</sup>	(s) pe-38, ie-1 and hrs	vDNA replication (unwinding the DNA by disrupting the hydrogen bounds between bases in the double helix, exposing the bases so that complementary DNA strands can be synthesized); shutdown of host protein synthesis; associated with the virogenic stroma	[144]
39K (E/L) <sup>C</sup>	(s) hrs and ie-1	Capable of binding to DNA; formation of virogenic stroma; involved in late promoter recognition and stabilization of late transcripts	[145]
p35 (E)	(s) hrs and ie-1	Anti-apoptotic gene that indirectly stimulates late transcription; binds and inhibits the action of caspases (enzyme mediating apoptosis)	[146]
ie-2 (E) <sup>A</sup>	(+) pe-38 (s) hrs and ie-1	Influence late transcription by acting on the transcription rate of early genes; cell cycle regulation and associated with the virogenic stroma	[146]
pe-38 (E)	(+) ie-1 and p143 (s) ie-2	Indirectly stimulates vDNA replication albeit not essential	[147]
ie-0 (E) <sup>A</sup>	(+) ie-1 (-) vp39	Infectious virus production	[148]
orf121 (E)	(+) ie-1	Early transactivator	[149]
lef-3 (E)		vDNA replication (ss)DNA binding protein which binds to ssDNA preventing the formation of secondary structures, renaturation of ssDNA and non-specific interactions between ssDNA and other proteins); associated with the virogenic stroma; mediates the transport of p143 to the nucleus for the binding to ssDNA	[150]
lef-1 (E)		vDNA replication (bind to ssDNA molecules allowing the DNA polymerase to recognize this "primer" and initiate the elongation of new DNA strands); DNA primase that interacts with lef-2	[151]
lef-2 (E)		vDNA replication (bind to ssDNA molecules allowing the DNA polymerase to recognize this "primer" and initiate the elongation of new DNA strands) and involved in very late transcription; DNA primase accessory factor that interacts with lef-1	[146]
dnapol (E)	(+) ie-1	Virus-encoded DNA polymerase essential for vDNA replication	[152]
DBP (E)		DNA binding protein capable of unwinding and annealing DNA; virogenic stroma production and structurally-normal nucleocapsid formation; maintenance of vDNA synthesis levels and stabilization of nascent vDNA by protecting mature viral genomes against nucleases at stages preceding packaging into virions	[153]

**Table 3 (cont.).** Index to named recombinant baculovirus gene/gene products.

Genes	Regulation	Function	References
lef-11 (E)		Involved in the process of vDNA replication (localizes in the nucleus); influence late transcription by acting on the transcription rate of early genes	[154]
lef-7 (E)		Involved in vDNA replication but not essential; influence late transcription by acting on the transcription rate of early genes	[155]
ME53 (E)		Stimulate gene viability; gene replication and nucleocapsid production albeit not essential for vDNA replication	[156]
pna (E)	(s) ie-1	Involved in vDNA replication but not essential	[146]
lef-4 (L)		Subunit of the RNA polymerase complex responsible for the mRNA capping by organizing the polymerase subunits	[144]
lef-8 (L)	Phosphorylated by pk-1	Subunits of the RNA polymerase complex required for late gene transcription	[157]
lef-9 (L)		Subunits of the RNA polymerase complex required for late gene transcription	[158]
p47 (L)		Subunits of the RNA polymerase complex required for late gene transcription	[147]
lef-5 (L)		Involved in late promoter recognition or stabilization of late transcripts; initiation factor rather than an elongation factor	[144]
lef-6 (L)		Involved in late promoter recognition or stabilization of late transcripts; influences the time for the onset of late transcription; transport of mRNA produced by very late genes	[159]
lef-10 (L)		Involved in late promoter recognition or stabilization of late transcripts	[147]
lef-12 (L)		Involved in late promoter recognition or stabilization of late transcripts	[142]
vlf-1 (LV/L)	(+) <i>polh</i> and <i>p10</i> (up to 10-fold)	Formation of nucleocapsids and their movement out of the virogenic stroma; protection of vDNA from DNases; vDNA processing although this is not clarified	[160]
ac66 (L)		Release of virions from nucleus (99% reduction in budded virus yields for ac66-null r-baculovirus)	[161]
vp91 (L)		Encodes a chitin binding domain essential for the movement of proteins and nutrients (localizes at the periphery of virogenic stroma)	[162]
p6.9 (LV/L) <sup>D</sup>	(+) vlf-1 (-) vlf-1	Originates DNA and p6.9 negatively charged by phosphorylation upon entry into infected cells (localizes to the nucleus matrix during infection)	[163]

Table 3 (cont.). Index to named recombinant baculovirus gene/gene products.

Genes	Regulation	Function	References
gp64 (E/L)	(s) ie-1	Form peplomers at the end of the baculoviruses; mediates virus binding to cell surface, probably through receptors such as phospholipids or proteins; efficient virion budding; escape of nucleocapsids from endosomes and entry into a wide variety of vertebrate cell lines; localizes in the plasma membrane at which the nucleocapsids will associate; formation of virus envelope	[164]
FP25 (L/VL)	(+) <i>polh</i>	Formation of nucleocapsids and formation of virogenic stroma	[165]
vp39 (L) <sup>E</sup>	(d) ie-0	Randomly distributed over the surface of the nucleocapsid; formation of nucleocapsids and virogenic stroma	[144]
gp16 (L)		Nucleocapsids envelope protein that is lost during transit through the cytoplasm and mediates the transport of nucleocapsids from the nucleus to the plasma membrane where they acquire a new envelope and bud from the cells	[166]
PP78/83 (L)	(+) actin polymerization	Capsid associated protein located at the end of nucleocapsids and adjacent to the <i>polh</i> gene being essential for virus infectivity	[167]
38K (L)	(+) early and late genes	Negatively regulates RNA polymerase by inhibiting RNA elongation thereby shutting off the transcription process	[167]
pk-1 (L/VL)	Phosphorylates Ief-8 and its inhibition impacts negatively on <i>polh</i> expression	Kinase involved in very late transcription	[168]
<i>p10</i> (VL)	(s) vif-1 Expression occurs earlier and at lower levels than <i>polh</i>	Interacts with tubulin forming microtubule-associated filaments or tube-like structures that surround the nucleus	[169]
<i>polh</i> (VL)	(s) vif-1, hr1 (up to 11-fold), pk-1, FP25 and ie-1 Expression later and at higher levels than <i>p10</i>	Not essential for virus replication or infection; protein PTM may be compromised due to its very late expression; requires a virus-encoded a amanitin RNA polymerase for its activation; may cause premature cellular disintegration	[170]
actin <sup>F</sup>		Transportation of naked nucleocapsids (250-300nm length and 30-60nm diameter) for the nucleus through nuclear pore complexes with channels of 38 nm	[171]

<sup>A</sup> the promoter strength is 10 to 50-fold lower than *polh* and *p10* gene promoters; <sup>B</sup> the promoter strength is 3.2 to 15-fold and <sup>12</sup> to 54-fold lower than *dnapol* and *polh* gene promoters, respectively; <sup>C</sup> the promoter strength is 5-fold lower than *polh* gene promoter; <sup>D</sup> the promoter strength is as efficient as the *polh* gene promoter; <sup>E</sup> the promoter strength is 2-fold lower than *polh* gene promoter; <sup>F</sup> Nuclear filamentous actin localized within the nucleus of host cells during the early stage of infection; (-) – down-regulates; (+) – up-regulates; (d) – down-regulated by; (s) – stimulated by; **DBP** – DNA binding protein; **dnapol** – DNA polymerase; **E** – activated in the early phase; **FP** – few polyhedra; **hrs** – homologous repeated sequences; **ie** – immediate early gene; **L** – activated in the late phase; **lef** – late expression factor; **orf** – open reading frame; **pk** – protein kinase; **polh** – polyhedrin; **PTM** – post-translational modifications; **ssDNA** – single-stranded DNA; **vdNA** – viral DNA; **VL** – activated in the very late phase; **vif** – very late factor

A percentage of the replicated vDNA is encapsidated in this phase, resulting in new virions, which then bud from the host cell. Progeny virus budding is known to begin around 16 hpi and endures for almost 36 h [43]; the newly budded viruses are able to start a second round of infection. The packaging of vDNA and the assembly of nucleocapsids is controlled by the very late factor (vlf)-1; its absence induces the formation of aberrant tube-like capsids. The way the virus senses that a genome sequence is complete and terminates the encapsidation process remains to be determined.

#### **Phase 4.** Very late gene expression (protein biosynthesis)

The transcription of the so-called “hyperexpressed” genes, such as those under the control of the *polh* or *p10* promoters, begins at the very late phase of the infection cycle, around 15 hpi (see genetic regulation at the *polh* node in Figure 4). Although regulated and stimulated by the same transcriptional factor (vlf-1), gene expression under the control of *p10* promoter is lower and occurs earlier than the *polh*. This difference may be related with the positive effect of the FP25 protein on the transcription rate of the *polh* promoter but not on the *p10* promoter.

#### **Phase 5.** Transcription shut-down

The shut-off of transcription may be related with the 38K gene which encodes a protein much similar to a set of enzymes related with the regulation of RNA polymerase.

#### **Phase 6.** The assembly process of RLPs

Newly synthesized proteins are degraded in lysosomes or enter the secretion pathway. Once the post-translational modifications are completed, the correctly folded proteins are released to the extracellular

matrix or assembled into RLPs. Details of the mechanisms involved in RLPs assembly are poorly understood; hardly any experimental information exists describing the assembly pathway, intra and inter-protein binding energies, rates and orders for assembly reactions, dissociation and association rates, and formation of stable structural subunits.

### 4.3. Comparison of selected models

Numerous comprehensive mathematical models have been developed for the baculovirus/insect cell system. A detailed description of how the most relevant variables of the system are addressed is presented below.

#### 4.3.1. Non-infected cell growth

The kinetics of uninfected cell population,  $n_u$ , is normally described using simple exponential growth models [43, 59, 125, 130, 132, 172-175]:

$$\frac{dn_u}{dt} = (\mu - k_d)n_u - v_{\text{inf}} \quad \text{Eq. 1.1}$$

where  $\mu$  is the cell specific growth rate and  $k_d$  the intrinsic cell death rate;  $v_{\text{inf}}$  represents the decrease in uninfected cell population due to viral infection (see section 4.3.3 for details).

Most correlations for the specific growth and death rate stem from a Monod-type equation with either the substrate [43, 107, 132, 175]

$$\mu = \mu_{\text{max}} \frac{S}{K_s + S} \quad \text{Eq. 1.2}$$

or the free surface area available for growth and the DO concentration

$$\mu(z, t) = \mu_{\text{max}} A(z, t) \left( \frac{C_L(z, t)}{K_{O_2} + C_L(z, t)} \right) \quad \text{Eq. 1.3}$$

as limiting factors [125, 173]. The maximum specific growth rate is  $\mu_{\max}$ ,  $z$  the vessel section ( $z=1$  for static operation mode),  $A(z,t)$  the dimensionless free area, which is dependent on the number of virus *per* cell,  $C_L$  the dissolved  $O_2$  concentration and  $K_{O_2}$  the half saturation constant.

#### 4.3.2. Binding and infection

The first mathematical models describing baculovirus infection were published by Licari and Bailey, Power and De Gooijer in 1992 (Table 4). These authors observed that the number of virus infecting an individual cell could be described by a Poisson distribution with the mean  $\lambda = \chi \cdot \text{MOI}$ :

$$P(t, nv) = \frac{\exp\left[-\chi \frac{v(t)}{n_u(t)}\right] \left[\chi \frac{v(t)}{n_u(t)}\right]^{nv}}{nv!} \quad \text{Eq.1.4}$$

where  $P(t, nv)$  is the probability of  $nv$  virus infecting a cell at time  $t$ ,  $\chi$  is a factor describing the effectiveness of infection (virus binding to the cell receptors and entry by endocytosis) and  $v$  is the virus concentration.

An important aspect in these models is the definition of the time for the onset of virus synthesis,  $\tau_v$ , which was considered to be a function of the number of viruses infecting a given cell population:

$$\tau_v(nv) = \tau_{v, \max} - \left( \frac{\tau_{v, \max} - \tau_{v, \min}}{nv_{\max} - 1} \right) (nv - 1) \quad \text{Eq. 1.5}$$

where  $\tau_{v, \min}$  and  $\tau_{v, \max}$  are the mathematical lower and upper bounds for the initiation of virus synthesis. Moreover, the virus synthesis was assumed to endure for  $\Delta\tau_v$  hours at a constant rate,  $k_v$ .



**Table 4.** Mathematical models describing the baculovirus infection process.

References	Virus attachment	Virus synthesis	Additional terms
Licari and Bailey (1992) and De Gooijer <i>et al.</i> (1992)	$\frac{dv}{dt} = -nv \times v_{inf}$	$+ \int_{\tau_v(mv)}^{\tau_v(mv)+\Delta\tau_v} A(1,t) k_v v_{inf}(t, \tau, nv) d\tau$	(A)
Power <i>et al.</i> (1992)	$\frac{dv}{dt} =$	$+ \int_{\tau_v(mv)}^{\tau_v(mv)+\Delta\tau_v} k_v v_{inf}(t, \tau, nv) d\tau$	
Power <i>et al.</i> (1994)	$\frac{dv}{dt} = -\alpha_b \times e^{(-\beta_b n_b)} n_b \times v$	$+ \int_{\tau_v(mv)}^{\tau_v(mv)+\Delta\tau_v} k_v v_{inf}(t, \tau, nv) d\tau$	$-k_{dv} \times v$ (B)
Power <i>et al.</i> (1996)	$\frac{dv}{dt} = \frac{1}{1 - e^{\left(-4\phi\pi n_b R \frac{kT}{6\pi\eta r}\right)}}$		
Kumar and Schuler (1995)	$\frac{dv}{dt} =$	$+k_r A(z,t) \frac{MOI(z,t)}{K_v/a + MOI(z,t)}$	
Dee and Shuler (1997)	$\frac{dv}{dt} = -k_a n_b v$		
Jang <i>et al.</i> (2001)	$\frac{dv}{dt} = -k_{ads}$	$+k_{prod}$	$-\xi \times v$ (C)
Petricevich <i>et al.</i> (2001)	$\frac{dv}{dt} =$	$+ \frac{DB \left(\frac{t}{C}\right)^{B-1}}{C} \left/ \left(1 + \left(\frac{t}{C}\right)^B\right)^2 \right.$	
Hu and Bentley (2000, 2001)	$\frac{dv}{dt} = -nv \times v_{inf}$	$+ \int_{\tau_v(mv)}^{\tau_v(mv)+\Delta\tau_v} k_v v_{inf}(t, \tau, nv) d\tau \times \left( \frac{\Delta\tau_v + \tau_v - \tau}{\Delta\tau_v} f(nv) \frac{S}{K_s + S} \right)$	$-P_{re-inf}(t, nv) \frac{\theta - \tau}{\theta} v_{inf}(t, \tau, nv)_{nv}$ (D)
Maranga <i>et al.</i> (2003)	$\frac{dv}{dt} = -k_a n_b v \times \phi$		
Gotoh <i>et al.</i> (2004)	$\frac{dv}{dt} = -nv \times v_{inf}$	$+ \int_{\tau_v(mv)}^{\tau_v(mv)+\Delta\tau_v} k_v v_{inf}(t, \tau, nv) d\tau \times f_{act}$	

<sup>A</sup> –  $A(1,t)=0$  in De Gooijer *et al.* (1992); <sup>B</sup> – virus decomposition term; <sup>C</sup> – physical term of the system; <sup>D</sup> – Re-infection term;  $dv/dt$  – extracellular concentration of baculoviruses per unit of time;  $\alpha_b$  and  $\beta_b$  – fitted parameters;  $n_b$  – number of uninfected cells plus cells infected for less than 15-20 h;  $\phi$  – proportion of successful collisions between virus particles and cells;  $R$  – host cell radius;  $k$  – Boltzmann's constant;  $T$  – temperature;  $\eta$  – viscosity of the medium;  $r$  – virus particle radius;  $k_r$  – infection rate constant;  $K_v$  – saturation-like parameter;  $a$  – surface area for growth per unit reactor volume;  $k_a$  – experimental measured attachment rate constant;  $\phi$  – efficiency factor (reflect the dependency of the binding rate upon the medium matrix composition);  $B$  – is a slope constant,  $C$  – is the time at which 50% of the response is reached;  $D$  – is the maximum bound virus;  $k_{ads}$  – virus depletion rate due to infection;  $k_{prod}$  – virus budding rate;  $\xi$  – the dependency of extracellular virus on the physical system;  $\theta$  – time windows for re-infection;  $S$  – rate-limiting substrate;  $K_s$  – Monod constant.

In 1994, Power and co-workers included in their model the concept of virus decomposition constant,  $k_{dv}$ ; virus attachment was also modified, being described by an empirical function dependent on virus concentration. In 1996, they used the Valentine and Allison model [176] to estimate the attachment of viruses to insect cells. Although neglecting the dependency between virus adsorption rate and virus concentration, results were consistent with the ones published previously.

Kumar and Schuler (1995) described virus formation using a Michaelis-Menten equation: at low virus concentration, virus synthesis follows a first-order kinetics as observed in static cultures [34] whereas at high concentration, a zero-order kinetics best describes the synthesis of new virions in suspension cultures [172]. In Petricevich *et al.* (2001), virus synthesis was modelled using the derivative of a sigmoidal curve [177].

The Dee and Schuler trafficking model is the most complete mathematical description of the baculovirus infection process [130, 178]. It includes receptor trafficking, binding, internalization, endosomal fusion and transport to nucleus. Virus production was not addressed in these studies.

In Jang *et al.* 2000 [136], an extra parameter  $\xi$  is used to account for the dependency of virus production on the physical system (culture volume, medium flux...). In batch cultures, this parameter is tuned to zero; in fed-batch cultures,  $\xi$  is a function of input and output fluxes.

None of the aforementioned deterministic models considered re-infection as a critical factor for progeny virus synthesis. Re-infection was only considered in Hu and Bentley's models [43, 107]. The re-infection rate was assumed to decrease linearly with time, consistent with experimental observations [178], and the probability of an infected cell being re-infected was described by a Poisson distribution function:

$$P_{re-inf}(t, nv) = \frac{\exp\left[-\chi_{re-inf} \frac{v(t)}{n_{inf}(t)}\right] \left[\chi_{re-inf} \frac{v(t)}{n_{inf}(t)}\right]^{nv}}{nv!} \quad \text{Eq. 1.6}$$

where  $\chi_{re-inf}$  is a factor describing the effectiveness of re-infection. The effect of infected cell's metabolic state, amount of internalized virus and nutrient's concentration on progeny virus synthesis was also incorporated in the model (Table 4).

In Gotoh *et al.* (2004), progeny virus synthesis is described by the Weibull distribution function instead of the traditional Poisson distribution:

$$P(t, \rho_i, \sigma_i) = \frac{\rho_i}{\sigma_i^{\rho_i}} \tau^{(\rho_i-1)} e^{\left[ -\left( \frac{\tau}{\sigma_i} \right)^{\rho_i} \right]} \quad \text{Eq. 1.7}$$

where  $\rho_i$  and  $\sigma_i$  are calculation parameters. To account for the cell density effect on virus synthesis rate, a dimensionless factor was also included:

$$f_{act} = \frac{X_{high} - X_{ICD}}{X_{high} - X_{low}} \quad \text{Eq. 1.8}$$

where  $X_{ICD}$  is the cell concentration at the  $t=0$ , and  $X_{low}$  and  $X_{high}$  are the lower and upper bound of uninfected cell concentration, respectively.

#### 4.3.3. Infected cell population

The equation describing the infected cell density,  $n_{inf}$ , consists of two terms: one for the increase in infected cells concentration due to binding of baculoviruses to uninfected cells and a second to represent the cell death. The increase in infected cells due to infection is generally described by:

$$\frac{dn_{inf}}{dt} = k_a n_u v \quad \text{Eq. 1.9}$$

where  $k_a$  is the attachment constant of virus to uninfected cells [174]. In Maranga *et al.* (2003), an extra parameter was added ( $\phi$ ) to cope with the dependency of the binding rate upon the medium matrix composition. On the other hand, Kumar and Schuler (1995) considered that the infection rate was saturable in virus concentration:

$$\frac{dn_{inf}(z,t)}{dt} = k_r n_u(z,t) \frac{MOI(z,t)}{K_v/a + MOI(z,t)} \quad \text{Eq. 1.10}$$

In Licari and Bailey (1992), Hu and Bentley (2000, 2001) and Gotoh *et al.* (2004), the infected cell population was described using a frequency function:

$$\frac{dn_{inf}}{dt} = P(t, nv) n_u \quad \text{Eq. 1.11}$$

where **P(t,nv)** is given by the Poisson or the Weibull distribution function.

The loss of infected cells due to infection can be represented by:

$$\frac{dn_{inf}}{dt} = -k_d n_{inf} \quad \text{Eq. 1.12}$$

where **k<sub>d</sub>** is the death rate of an infected cell. The **k<sub>d</sub>** was assumed constant throughout the infection process in Gotoh *et al.* (2004). In contrast, Power *et al.* (1992) and Hu and Bentley (2000, 2001) described **k<sub>d</sub>** as a function of the time a cell has been infected. This implies that **k<sub>d</sub>** is constant at **k<sub>d1</sub>** from time of infection until the time when rapid death of the population has been observed experimentally (**t<sub>d</sub>**); upon this threshold, the rate increases significantly (**k<sub>d2</sub>**). The **t<sub>d</sub>** has been described as a constant [174] or as dependent on the MOI [43, 107].

#### 4.3.4. vDNA and mRNA

The intracellular dynamics of vDNA was defined by Jang *et al.* (2000) as:

$$\frac{ddna}{dt} = r_{dna,inf} + r_{dna,rep} - r_{bud} - dna \times \xi \quad \text{Eq. 1.13}$$

where **r<sub>dna,inf</sub>** represents the rate of infection (i.e. number of viruses penetrating uninfected and infected cells within a specific time interval), **r<sub>dna,rep</sub>** the vDNA replication rate, **r<sub>bud</sub>** the virus budding rate and **ξ** a parameter accounting for the loss of vDNA templates due to the physical

system. Remaining models describing the replication of vDNA neglect the influence of virus budding on the levels of intracellular vDNA [68, 179].

The dynamics of vDNA transcription into mRNA was defined as:

$$\frac{drna}{dt} = r_{rna,s} - rna \times \xi \quad \text{Eq. 1.14}$$

where  $r_{rna,s}$  is the formation rate of mRNA [136]. In both kinetics, vDNA replication and transcription, Jang and co-workers assumed that all the rates were constant throughout the process.

#### 4.3.5. Protein synthesis

The net synthesis of heterologous protein can be defined by:

$$\frac{dprot}{dt} = r_{prot,s} - r_{prot,d} - r_{prot,sec} - prot \times \xi \quad \text{Eq. 1.15}$$

where  $r_{prot,s}$  is the protein synthesis rate,  $r_{prot,d}$  the protein degradation rate and  $r_{prot,sec}$  the protein secretion rate.

Most models describe the protein synthesis rate as a constant [42, 136, 173-175, 177, 180-182]; others explicitly describe it as a function of process-related parameters:

$$r_{prot,s} = \frac{\zeta}{\tau_L(nv) - \tau_p(nv)} A(1, t) \quad \text{Ref.: [125]} \quad \text{Eq. 1.16}$$

$$r_{prot,s} = k_{prot,s} \frac{\Delta\tau_p + \tau_p - \tau}{\Delta\tau_p} \frac{S}{K_s + S} f(nv) \quad \text{Ref.: [43, 107]} \quad \text{Eq. 1.17}$$

$$r_{prot,s} = k_{prot,s} P(t, nv) n_u \quad \text{Ref.: [132]} \quad \text{Eq. 1.18}$$

where  $\zeta$  is the maximum quantity of foreign protein produced per cell,  $\tau_L$  and  $\tau_p$  the times post-infection for the onset of cell lysis and protein

synthesis, respectively (which are a function of the number of viruses infecting a given cell population – similar to  $\tau_v$  - Eq. 1.6),  $k_{\text{prot},s}$  the protein synthesis constant and  $\Delta\tau_p$  the time-window over which recombinant protein is produced.

The protein degradation rate is written as  $r_{\text{prot},d} = k_{\text{prot},d} \times \text{prot}$ , where  $k_{\text{prot},d}$  is the degradation constant [132]. The quantity of protein secreted to the extracellular medium is proportional to the concentration of intracellular protein and constrained in time  $[\tau_{LK}, \tau_L]$  ( $\tau_{LK}$  is the time post-infection at which leakage begins); hence,  $r_{\text{prot},sec} = k_{\text{prot},sec} \times \text{prot}$ , where  $k_{\text{prot},sec}$  is the protein secretion constant [125].

#### 4.3.6. The assembly of VLPs

The assembly of complex protein structures composed by several protein subunits is poorly understood. Little experimental information exists describing intra and inter-subunit binding energies, rates and orders for assembly reactions, and formation of nucleating structures.

So far, there are two models dealing with VLP formation in the baculovirus/insect cell system [43, 59, 107]. In Maranga *et al.* (2003), a mathematical model was developed to account for the effects of virus multiplicity, cell concentration and nutrient limitations on VLP productivity; no detailed description of the assembly process is formulated. The Hu and Bentley model is the only one that tries to describe the assembly of VLPs; it uses a thermodynamic equilibrium based approach where the Gibbs free energy of subunit association and concentration of viral protein subunits are key factors for the efficiency of the overall process.

Nevertheless, the mathematical modelling of VLP formation in the baculovirus/insect cell system is not a closed chapter, especially in what

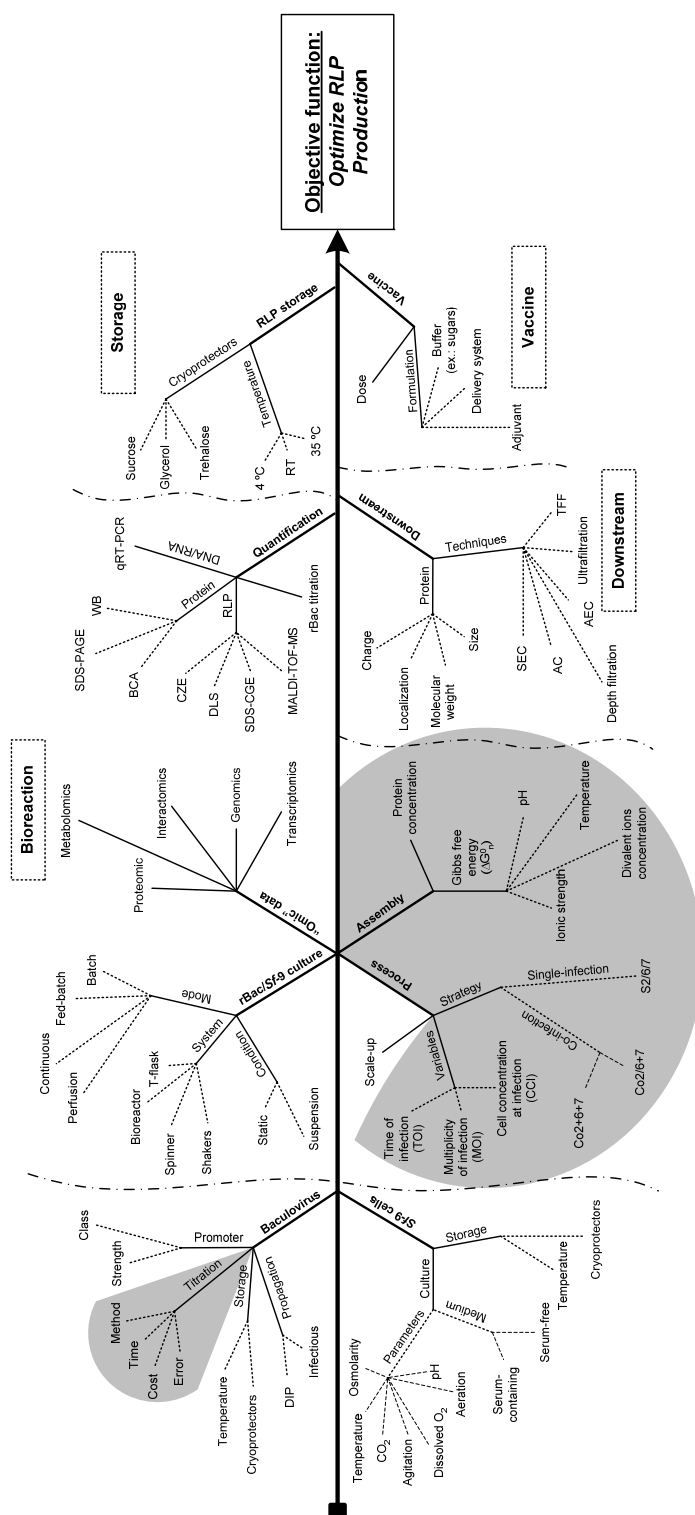
concerns post-infection intracellular dynamics. Probably due to the lack of knowledge and experimental evidence on viral genes expression regulation, there are no structured dynamic models in the literature that describe early, late and very late gene expression cascade upon infection. Moreover, VLP formation is ruled by thermodynamic and kinetic factors. While a few studies addressed the assembly of simple structures, the complexity of multi-layered particles has never been attempted before.

## **5. SCOPE OF THE THESIS**

The secret to overcoming the bioengineering challenges imposed by the production of RLPs in the baculovirus/insect cell system is to define properly the main steps involved in the overall process, identify the main factors affecting each individual step and delineate optimization schemes capable of maximizing the objective function – RLP production (Figure 5).

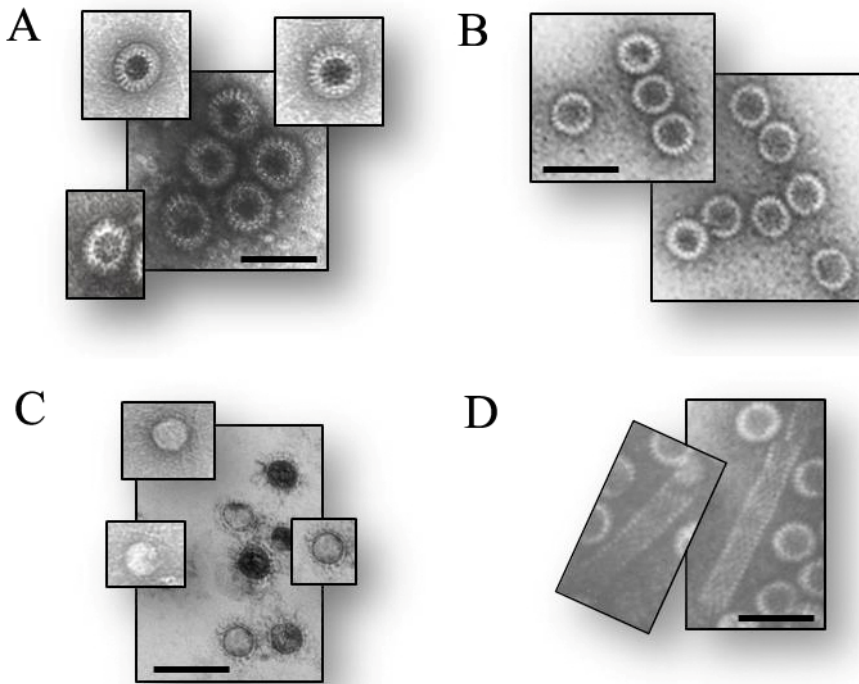
The diversity of protein structures observed at the end of RLP production runs indicates that the assembly process is highly inefficient (Figure 6). Contaminants such as trimers of vp6 and vp7, vp6 tube-like structures, single-layered vp2 particles, double layered particles of vp2 and vp6 (DLPs) or RLPs lacking one or more subunits represent almost 88% of the total mass of proteins expressed.

To understand the entirety of the processes happening during the production of a RLP, from the genetics of the viral vector to the infection strategy, titer and structure of assembled particles, a bottom-up Systems Biology approach was pursued.



**Figure 5.** Ishikawa diagram of the main bioengineering challenges of RLP production in the baculovirus/insect cell system. Highlighted in grey are the issues addressed in this Thesis. Abbreviations: **DIP** – defective interfering particles; **qRT-PCR** – real time quantitative polymerase chain reaction; **WB** – Western blot; **SDS-PAGE** – sodium dodecyl sulfate polyacrylamide gel electrophoresis; **BCA** – Bicinchronic acid assay; **CZE** – capillary zone electrophoresis; **DLS** – dynamic light scattering; **SDS-CGE** – sodium dodecyl sulphate capillary gel electrophoresis; **MALDI-TOF-MS** – Matrix-assisted laser-desorption/ionization - Time-of-flight - Mass spectrometry; rBac – recombinant baculovirus; SEC – size exclusion chromatography; AC – affinity chromatography; AEC – anion exchange chromatography; **TFF** – tangential flow filtration





**Figure 6.** Electron micrograph of negatively stained RLPs (A), DLPs (B), single-layered vp2 particles (C) and vp6 tube-like structures (D) [81]. Scale bars correspond to 100 nm.

A basic requirement for data generation, modelling and optimisation of the baculovirus/insect cell system is measuring and controlling the MOI, which is the most important manipulated variable in this process. Unfortunately, measuring baculovirus titers is costly and inaccurate while the ability to control MOI to sufficient accuracy is still debatable. As such, this Thesis starts (Chapter 2) with a comparative study of titration methods in what concerns accuracy, cost and work intensity. The most critical issues of each method were identified and suggestions of improvements were presented.

Mathematical models for *in silico* baculovirus/insect cell simulation and prediction of protein synthesis and self assembly of polyhedral protein complexes were developed. The present thesis integrates a complex

network of mathematical/statistical model formulations (stochasticity, segregation, molecular scale, kinetics and thermodynamics), with experimental intra- and extra-cellular data (vDNA replication, very late transcription, protein synthesis and RLP formation). It addresses the most relevant mechanisms of the baculovirus/insect cell system for recombinant protein production and self assembly of proteins, namely:

- Stochastic vs deterministic infection: the attachment of baculoviruses to insect cell's membrane was described as a stochastic process. This allowed studying the differences between deterministic and stochastic infection and the implications of stochastic infection on protein expression at high and low MOIs (Chapter 3);
- Segregated infected cells: due to random baculovirus entry into insect cells, each infected cell undergoes distinctive infection dynamics with impact on viral proteins expression and RLP syntethesis. To take this effect into account, population models were developed (Chapter 3);
- Thermodynamic RLP assembly: a thermodynamic equilibrium-based model describing the assembly process of RLPs. A detailed *in silico* analysis of the biophysical mechanisms involved in the assembly process and their controllability is presented (Chapter 4);
- Infection strategy: a deterministic structured model was developed to analyze the relevant intracellular events (baculovirus trafficking, vDNA replication, mRNA synthesis, protein expression and RLP assembly) and their variability with the MOI at different infection strategies (single- and co-infection) (Chapter 5);

*In silico* analysis of the baculovirus/insect cell system was performed in what concerns the sensitivity of RLPs and constitutive proteins concentration in relation to model parameters (Chapters 6) and to process

degrees of freedom such as the MOI or the Gibbs free energy of protein subunit association (Chapters 3 to 6). It should be noted that the latter parameter can be considered as an indirect process degree of freedom as it can be manipulated to some extent by designing the environment for protein assembly.

The theoretical effect of MOI and promoter strength on RLP production, at single- and co-infection strategies, was explored using the *in silico* stochastic model developed in Chapter 3. Optimal culture conditions for RLP production were determined and potentially better infection strategies, capable of increasing RLP yields, were proposed (Chapters 6).

Summarizing, this thesis provides a set of computational models of the baculovirus/insect cell system to assist the development of novel and improved optimization schemes for production of virus particles, viral proteins and/or complex protein structures such as virus-like particles.

## 6. REFERENCES

1. O'Reilly, D.R., Miller, A.D., and Luckow, V.A., **1994**. Baculovirus expression vectors: A laboratory manual. Oxford University Press. New York, USA.
2. Summers, M.D., Smith, G.E., **1987**. A manual of methods for baculovirus vectors and insect cell culture procedures. Texas agricultural Experimental Station.
3. Jiang, B., *et al.*, **1999**. Heterotypic protection from rotavirus infection in mice vaccinated with virus-like particles. *Vaccine* 7-8 (17): 1005-13.
4. Kang, S.M., *et al.*, **2009**. Influenza vaccines based on virus-like particles. *Virus Res*
5. Davis, T.R. and Wood, H.A., **1995**. Intrinsic glycosylation potentials of insect cell cultures and insect larvae. *In Vitro Cell Dev Biol Anim* 9 (31): 659-63.
6. Ahn, M.-H., *et al.*, **2008**. Production of therapeutic proteins with baculovirus expression system in insect cell. *Entomological Research* S1 (38): 71-78.
7. Gao, G., *et al.*, **2003**. High throughput creation of recombinant adenovirus vectors by direct cloning, green-white selection and I-Sce I-mediated rescue of circular adenovirus plasmids in 293 cells. *Gene Ther* 22 (10): 1926-30.
8. Parry, S., Koi, H., and Strauss, J.F., **1999**. Transplacental drug delivery: gene and virus delivery to the trophoblast. *Adv Drug Deliv Rev* 1 (38): 69-80.
9. Fromentin, R., *et al.*, **2007**. A method for in vitro assembly of hepatitis C virus core protein and for screening of inhibitors. *Anal Biochem* 1 (366): 37-45.
10. Lapid, B., *et al.*, **2009**. Innovative applications for insect viruses: towards insecticide sensitization. *Trends Biotechnol* 4 (27): 190-8.
11. Artenstein, A.W. and Grabenstein, J.D., **2008**. Smallpox vaccines for biodefense: need and feasibility. *Expert Rev Vaccines* 8 (7): 1225-37.

12. Fischlechner, M. and Donath, E., **2007**. Viruses as building blocks for materials and devices. *Angew Chem Int Ed Engl* 18 (46): 3184-93.
13. Airenne, K.J., *et al.*, **2009**. Baculovirus-Mediated Gene Transfer: An Emerging Universal Concept. *Gene and Cell Therapy: Therapeutic mechanisms and strategies*, third edition, ed. Templeton, N.S. CRC Press. 263-291. Boca Raton, FL.
14. Ghosh, S., *et al.*, **2002**. Baculovirus as mammalian cell expression vector for gene therapy: an emerging strategy. *Mol Ther* 1 (6): 5-11.
15. Boyce, F.M. and Bucher, N.L., **1996**. Baculovirus-mediated gene transfer into mammalian cells. *Proc Natl Acad Sci U S A* 6 (93): 2348-52.
16. Singh, P., Gonzalez, M.J., and Manchester, M., **2006**. Viruses and their uses in nanotechnology. *Drug development research* 1 (67): 23-41.
17. Summers, M.D. and Anderson, D.L., **1972**. Granulosis virus deoxyribonucleic acid: a closed, double-stranded molecule. *J Virol* 4 (9): 710-713.
18. Burgess, S., **1977**. Molecular Weights of Lepidopteran Baculovirus DNAs: Derivation by Electron Microscopy. *J Gen Virol* (37): 501-510.
19. Tweeten, K.A., Bulla, L.A., and Consigli, R.A., **1980**. Characterization of an Extremely Basic Protein Derived from Granulosis Virus Nucleocapsids. *J Virol* 2 (33): 866-876.
20. Thiem, S.M. and Miller, L.K., **1989**. Identification, sequence, and transcriptional mapping of the major capsid protein gene of the baculovirus *Autographa californica* nuclear polyhedrosis virus. *J Virol* 5 (63): 2008-2018.
21. Fraser, M.J., **1986**. Ultrastructural observations of virion maturation in *Autographa californica* nuclear polyhedrosis virus infected *Spodoptera frugiperda* cell cultures. *Journal of ultrastructure and molecular structure research* 1-3 (95): 189-195.
22. Belyaev, A.S. and Roy, P., **1993**. Development of baculovirus triple and quadruple expression vectors: co-expression of three or four bluetongue virus proteins and the synthesis of bluetongue virus-like particles in insect cells. *Nucleic Acids Res* 5 (21): 1219-23.
23. Jorio, H., Tran, R., and Kamen, A., **2006**. Stability of Serum-Free and Purified Baculovirus Stocks under Various Storage Conditions. *Biotechnol. Prog.* (22): 319-325.
24. Hu, W.S. and Oberg, M.G., **1990**. Monitoring and control of animal cell reactors: Biochemical Engineering considerations. *Large-Scale Mammalian Cell Culture Technology*, ed. Lubiniecki, A.S. Marcel Dekker, Inc. 451-482. New York.
25. Murhammer, D.W. and Gooch, C.F., **1990**. Sparged animal cell bioreactors: mechanism of cell damage and Pluronic F-68 protection. *Biotechnol Prog* 5 (6): 391-7.
26. Murhammer, D. and Gooch, C., **1988**. Scale up of insect cell cultures: protective effects of Pluronic F-68. *Biotechnol* (6): 1411-1418.
27. Elias, C.B., *et al.*, **2000**. Enhanced growth of Sf-9 cells to a maximum density of  $5.2 \times 10^7$  cells per mL and production of beta-galactosidase at high cell density by fed batch culture. *Biotechnol Bioeng* 4 (68): 381-8.
28. Doekel, S., Eppelmann, K., and Marahiel, M.A., **2002**. Heterologous expression of nonribosomal peptide synthetases in *B. subtilis*: construction of a bi-functional *B. subtilis*/E. coli shuttle vector system. *FEMS Microbiol Lett* 2 (216): 185-91.
29. Repaske, R. and Clayton, M.A., **1978**. Control of *Escherichia coli* growth by CO<sub>2</sub>. *J Bacteriol* 3 (135): 1162-4.
30. Konz, J.O., King, J., and Cooney, C.L., **1998**. Effects of oxygen on recombinant protein expression. *Biotechnol Prog* 3 (14): 393-409.
31. Cruz, P.E., *et al.*, **1998**. Optimization of the production of virus-like particles in insect cells. *Biotechnol Bioeng* 4 (60): 408-18.
32. Morris, T.D. and Miller, L.K., **1992**. Promoter influence on baculovirus-mediated gene expression in permissive and nonpermissive insect cell lines. *J Virol* 12 (66): 7397-405.
33. Wickham, T.J., *et al.*, **1995**. Comparison of Different Cell Lines for the Production of Recombinant Baculovirus Proteins. *Baculovirus Expression Protocols*, ed. Richardson, C.D. Humana Press. Vol. 39. 385-395.
34. Wickham, T.J., *et al.*, **1992**. Screening of insect cell lines for the production of recombinant proteins and infectious virus in the baculovirus expression system. *Biotechnol Prog* 5 (8): 391-6.
35. Caron, A.W., Archambault, J., and Massie, B., **1990**. High-level recombinant protein production in bioreactors using the baculovirus-insect cell expression system. *Biotechnol Bioeng* 11 (36): 1133-40.
36. Luckow, V.A. and Summers, M.D., **1988**. Signals important for high-level expression of foreign genes in *Autographa californica* nuclear polyhedrosis virus expression vectors. *Virology* 1 (167): 56-71.
37. Chai, H., *et al.*, **1996**. Insect cell line dependent gene expression of recombinant human necrosis factor. *Enzyme Microb. Technol.* (18): 126-132.

38. Nwe, N., *et al.*, **2006**. Expression of hemagglutinin protein from the avian influenza virus H5N1 in a baculovirus/insect cell system significantly enhanced by suspension culture. *BMC Microbiol* (6): 16.
39. Matsuura, Y., *et al.*, **1988**. Purification and characterization of IL-6 (IFN-beta 2/BSF-2) expressed in insect cells by a baculo virus vector. *J. Interferon Res.* (8): Suppl.1, S151.
40. Chiou, C.J. and Wu, M.C., **1990**. Expression of human granulocyte-macrophage colony-stimulating factor gene in insect cells by a baculovirus vector. *FEBS Lett* 2 (259): 249-53.
41. Strauss, A., *et al.*, **2007**. Improved expression of kinases in Baculovirus-infected insect cells upon addition of specific kinase inhibitors to the culture helpful for structural studies. *Protein Expr Purif* 2 (56): 167-76.
42. Tsao, E.I., *et al.*, **1996**. Production of parvovirus B19 vaccine in insect cells co-infected with double baculoviruses. *Biotechnol Bioeng* 2 (49): 130-138.
43. Hu, Y.-C. and Bentley, W.E., **2000**. A kinetic and statistical-thermodynamic model for baculovirus infection and virus-like particle assembly in suspended insect cells. *Chem Eng Sci* (55): 3991-4008.
44. Hink, W.F., *et al.*, **1991**. Expression of three recombinant proteins using baculovirus vectors in 23 insect cell lines. *Biotechnol. Prog.* (7): 9-14.
45. Lee, G.Y., *et al.*, **2006**. Expression and characterization of human vascular endothelial growth factor (VEGF165) in insect cells. *Protein Expr Purif* 2 (46): 503-9.
46. Nguyen, B., *et al.*, **1993**. Fed-batch culture of insect cells: a method to increase the yield of recombinant human nerve growth factor (rhNGF) in the baculovirus expression system. *J Biotechnol* 2 (31): 205-17.
47. Rosenblatt, J., *et al.*, **1993**. Purification and crystallization of human cyclin-dependent kinase 2. *J Mol Biol* 4 (230): 1317-9.
48. Wojchowslu, D.M., Lorkin, S.H., and Sytkowski, A.J., **1987**. Active human erythropoietin expressed in insect cells using a baculovirus vector: a role for N-linked oligosaccharide. *Biochim. Biophys. Acta* (910): 224-232.
49. Whang, Y., *et al.*, **1987**. Expression of the Epstein-Barr virus gp350/220 gene in rodent and primate cells. *J Virol* 6 (61): 1796-807.
50. Bonafe, N., *et al.*, **2009**. A recombinant West Nile virus envelope protein vaccine candidate produced in *Spodoptera frugiperda* expresSF+ cells. *Vaccine* 2 (27): 213-22.
51. Tokarski, J.S., *et al.*, **2006**. The structure of Dasatinib (BMS-354825) bound to activated ABL kinase domain elucidates its inhibitory activity against imatinib-resistant ABL mutants. *Cancer Res* 11 (66): 5790-7.
52. Martinez-Torrecuadrada, J.L., *et al.*, **2005**. An efficient expression system for the production of functionally active human LKB1. *J Biotechnol* 1 (115): 23-34.
53. Taylor, A.L., *et al.*, **2006**. High yield of biologically active recombinant human amelogenin using the baculovirus expression system. *Protein Expr Purif* 1 (45): 43-53.
54. Jiang, B., *et al.*, **1998**. Synthesis of rotavirus-like particles in insect cells: comparative and quantitative analysis. *Biotechnol Bioeng* 3 (60): 369-74.
55. Bonnet, S., *et al.*, **2006**. Soluble and glyco-lipid modified baculovirus Plasmodium falciparum C-terminal merozoite surface protein 1, two forms of a leading malaria vaccine candidate. *Vaccine* 33-34 (24): 5997-6008.
56. Takahashi, S., *et al.*, **2007**. High-level expression of recombinant active human renin in Sf-9 cells: rapid purification and characterization. *Biosci Biotechnol Biochem* 10 (71): 2610-3.
57. Ingley, E., *et al.*, **1991**. Production and purification of recombinant human interleukin-5 from yeast and baculovirus expression systems. *Eur J Biochem* 3 (196): 623-9.
58. Reuveny, S., *et al.*, **1993**. Effect of temperature and oxygen on cell growth and recombinant protein production in insect cell cultures. *Appl Microbiol Biotechnol* 5 (38): 619-23.
59. Maranga, L., Brazao, T.F., and Carrondo, M.J., **2003**. Virus-like particle production at low multiplicities of infection with the baculovirus insect cell system. *Biotechnol Bioeng* 2 (84): 245-53.
60. Steiner, H., *et al.*, **1988**. Human tissue-type plasminogen activator synthesized by using a baculovirus vector in insect cells compared with human plasminogen activator produced in mouse cells. *Gene* 2 (73): 449-57.
61. Kuroda, K., *et al.*, **1990**. The oligosaccharides of influenza virus hemagglutinin expressed in insect cells by a baculovirus vector. *Virology* 2 (174): 418-29.
62. Mallick, S., *et al.*, **1992**. Expression and purification of human granulocyte colony stimulating factor using a eukaryotic vector expression system. *FASEB J.* 1 (6): A48.
63. Mena, J.A., Ramirez, O.T., and Palomares, L.A., **2005**. Quantification of rotavirus-like particles by gel permeation chromatography. *J Chromatogr B Analyt Technol Biomed Life Sci* 1-2 (824): 267-76.

64. Davidson, D.J., Fraser, M.J., and Castellino, F.J., **1990**. Oligosaccharide processing in the expression of human plasminogen cDNA by lepidopteran insect (*Spodoptera frugiperda*) cells. *Biochemistry* 23 (29): 5584-90.
65. Chazenbalk, G.D. and Rapoport, B., **1995**. Expression of the extracellular domain of the thyrotropin receptor in the baculovirus system using a promoter active earlier than the polyhedrin promoter. Implications for the expression of functional highly glycosylated proteins. *J Biol Chem* 4 (270): 1543-9.
66. Roldao, A., *et al.*, **2009**. Error assessment in recombinant baculovirus titration: evaluation of different methods. *J Virol Methods* 1 (159): 69-80.
67. Kärkkäinen, H.-R., *et al.*, **2009**. A 96-well format for a high-throughput baculovirus generation, fast titrating and recombinant protein production. Unpublished results.
68. Vieira, H., *et al.*, **2005**. Triple layered rotavirus VLP assembly: kinetics of vector replication, mRNA stability and recombinant protein production. *J Biotechnol*, 1 (120): 72-82.
69. Bernal, V., *et al.*, **2009**. Cell density effect in the baculovirus-insect cells system: a quantitative analysis of energetic metabolism. *Biotechnol Bioeng* 1 (104): 162-80.
70. Carinhas, N., *et al.*, **2009**. Baculovirus production for gene therapy: the role of cell density, multiplicity of infection and medium exchange. *Appl Microbiol Biotechnol* 6 (81): 1041-1049.
71. Kompier, R., Tramper, J., and Vlak, J.M., **1988**. A continuous process for the production of baculovirus using insect cell cultures. *Biotechnol. Lett.* (10): 849-854.
72. Caron, A.W., *et al.*, **1994**. Baculovirus expression system scaleup by perfusion of high-density Sf-9 cell cultures. *Biotechnol Bioeng* 9 (43): 881-91.
73. Roldão, A., *et al.*, **2008**. Stochastic simulation of protein expression in the baculovirus/insect cells system. *Computers & Chemical Engineering* 1-2 (32): 68-77.
74. Ferrance, J.P., Goel, A., and Ataai, M.M., **1993**. Utilization of glucose and amino acids in insect cell cultures: Quantifying the metabolic flows within the primary pathways and medium development. *Biotechnol Bioeng* 6 (42): 697-707.
75. Bedard, C., Tom, R., and Kamen, A., **1993**. Growth, nutrient consumption, and end-product accumulation in Sf-9 and BTI-EAA insect cell cultures: insights into growth limitation and metabolism. *Biotechnol Prog* 6 (9): 615-24.
76. Goodwin, R.H., **1991**. Replacement of vertebrate serum with lipids and other factors in the culture of invertebrate cells, tissues, parasites, and pathogens. *In Vitro Cell Dev Biol* 6 (27A): 470-8.
77. Mitsuhashi, J., **1989**. Nutritional requirements of insect cells in vitro. *Invertebrate Cell System Applications*, ed. Mitsuhashi, J. CRC Press. 3-20.
78. Wang, M.Y., Kwong, S., and Bentley, W.E., **1993**. Effects of oxygen/glucose/glutamine feeding on insect cell baculovirus protein expression: a study on epoxide hydrolase production. *Biotechnol Prog* 4 (9): 355-61.
79. Mendonca, R.Z., Palomares, L.A., and Ramirez, O.T., **1999**. An insight into insect cell metabolism through selective nutrient manipulation. *J Biotechnol* (72): 61-75.
80. Palomares, L.A. and Ramirez, O.T., **1996**. The effect of dissolved oxygen tension and the utility of oxygen uptake rate in insect cell culture. *Cytotechnology* (1-3) (22 ): 225-237.
81. Mellado, M.C., *et al.*, **2008**. Sodium dodecyl sulfate-capillary gel electrophoresis analysis of rotavirus-like particles. *J Chromatogr A* 1 (1192): 166-72.
82. Franco, C.F., *et al.*, **2009**. Monitoring virus-like particle and viral protein production by intact cell MALDI-TOF mass spectrometry. *Talanta* In press.
83. Mellado, M.C.M., *et al.*, **2009**. Assessment of RLPs production and assembly efficiency using capillary zone electrophoresis. In manuscript.
84. Cruz, P.E., *et al.*, **1999**. Production and quality analysis of Pr55gag particles produced in baculovirus-infected insect cells. *Journal of Chemical Technology & Biotechnology* 2 (72): 149-158.
85. Guerrero, R.A., *et al.*, **2001**. Recombinant Norwalk virus-like particles administered intranasally to mice induce systemic and mucosal (fecal and vaginal) immune responses. *J Virol* 20 (75): 9713-22.
86. Ball, J.M., *et al.*, **1999**. Recombinant Norwalk virus-like particles given orally to volunteers: phase I study. *Gastroenterology* 1 (117): 40-8.
87. Ruffing, M., Zentgraf, H., and Kleinschmidt, J.A., **1992**. Assembly of viruslike particles by recombinant structural proteins of adeno-associated virus type 2 in insect cells. *J Virol* 12 (66): 6922-30.
88. Sun, Y., *et al.*, **2009**. Protection against lethal challenge by Ebola virus-like particles produced in insect cells. *Virology* 1 (383): 12-21.
89. Chung, Y.C., *et al.*, **2006**. Expression, purification and characterization of enterovirus-71 virus-like particles. *World J Gastroenterol* 6 (12): 921-7.

90. Matassov, D., Cupo, A., and Galarza, J.M., **2007**. A novel intranasal virus-like particle (VLP) vaccine designed to protect against the pandemic 1918 influenza A virus (H1N1). *Viral Immunol* 3 (20): 441-52.
91. D'Aoust, M.A., *et al.*, **2008**. Influenza virus-like particles produced by transient expression in *Nicotiana benthamiana* induce a protective immune response against a lethal viral challenge in mice. *Plant Biotechnol J* 9 (6): 930-40.
92. Latham, T. and Galarza, J.M., **2001**. Formation of wild-type and chimeric influenza virus-like particles following simultaneous expression of only four structural proteins. *J Virol* 13 (75): 6154-65.
93. Mahmood, K., *et al.*, **2008**. H5N1 VLP vaccine induced protection in ferrets against lethal challenge with highly pathogenic H5N1 influenza viruses. *Vaccine* 26 (26): 5393-9.
94. Keating, G.M. and Noble, S., **2003**. Recombinant hepatitis B vaccine (Engerix-B): a review of its immunogenicity and protective efficacy against hepatitis B. *Drugs* 10 (63): 1021-1051.
95. Venters, C., Graham, W., and Cassidy, W., **2004**. Recombivax-HB: perspectives past, present and future. *Expert Rev Vaccines* 2 (3): 119-129.
96. Seo, H.S., *et al.*, **2008**. Analysis and characterization of hepatitis B vaccine particles synthesized from *Hansenula polymorpha*. *Vaccine* 26 (26): 4138-4144.
97. Madalinski, K., *et al.*, **2001**. Antibody responses to preS components after immunization of children with low doses of BioHepB. *Vaccine* 19 (20): 92-7.
98. Jeong, S.H., *et al.*, **2004**. Immunization with hepatitis C virus-like particles induces humoral and cellular immune responses in nonhuman primates. *J Virol* 13 (78): 6995-7003.
99. Lechmann, M., *et al.*, **2001**. Hepatitis C virus-like particles induce virus-specific humoral and cellular immune responses in mice. *Hepatology* 2 (34): 417-23.
100. Gheysen, D., *et al.*, **1989**. Assembly and release of HIV-1 precursor Pr55gag virus-like particles from recombinant baculovirus-infected insect cells. *Cell* 1 (59): 103-112.
101. Shioda, T. and Shibuta, H., **1990**. Production of human immunodeficiency virus (HIV)-like particles from cells infected with recombinant vaccinia viruses carrying the gag gene of HIV. *Virology* 1 (175): 139-48.
102. Sakuragi, S., *et al.*, **2002**. HIV type 1 Gag virus-like particle budding from spheroplasts of *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* 12 (99): 7956-61.
103. Einstein, M.H., *et al.*, **2009**. Comparison of the immunogenicity and safety of Cervarix(R) and Gardasil(R) human papillomavirus (HPV) cervical cancer vaccines in healthy women aged 18-45 years. *Hum Vaccin* 10 (5): 505-515.
104. Lai, W.B. and Middelberg, A.P.J., **2002**. The production of human papillomavirus type 16 vaccine product from *Escherichia coli* inclusion bodies. *Bioprocess Biosyst Eng* 25 (2): 121-128.
105. Zhou, J., *et al.*, **1991**. Expression of vaccinia recombinant HPV 16 L1 and L2 ORF proteins in epithelial cells is sufficient for assembly of HPV virion-like particles. *Virology* 1 (185): 251-7.
106. Heegaard, E.D. and Brown, K.E., **2002**. Human parvovirus B19. *Clin Microbiol Rev* 3 (15): 485-505.
107. Hu, Y.C. and Bentley, W.E., **2001**. Effect of MOI ratio on the composition and yield of chimeric infectious bursal disease virus-like particles by baculovirus co-infection: deterministic predictions and experimental results. *Biotechniques* 1 (75): 104-119.
108. Swenson, D.L., *et al.*, **2004**. Generation of Marburg virus-like particles by co-expression of glycoprotein and matrix protein. *FEMS Immunol Med Microbiol* 1 (40): 27-31.
109. Huang, Z., *et al.*, **2005**. Virus-like particle expression and assembly in plants: hepatitis B and Norwalk viruses. *Vaccine* 15 (23): 1851-8.
110. Brautigam, S., Snezhkov, E., and Bishop, D.H., **1993**. Formation of poliovirus-like particles by recombinant baculoviruses expressing the individual VP0, VP3, and VP1 proteins by comparison to particles derived from the expressed poliovirus polyprotein. *Virology* 2 (192): 512-24.
111. Martinez, C., *et al.*, **1992**. Production of porcine parvovirus empty capsids with high immunogenic activity. *Vaccine* 10 (10): 684-90.
112. Yu, F., *et al.*, **2001**. Characterization of Rous sarcoma virus Gag particles assembled in vitro. *J Virol* 6 (75): 2753-64.
113. Yamshchikov, G.V., *et al.*, **1995**. Assembly of SIV virus-like particles containing envelope proteins using a baculovirus expression system. *Virology* 1 (214): 50-8.
114. Yang, C., Ye, L., and Compans, R.W., **2008**. Protection against filovirus infection: virus-like particle vaccines. *Expert Rev Vaccines* 3 (7): 333-44.
115. Wagner, R., *et al.*, **1996**. Safety and immunogenicity of recombinant human immunodeficiency virus-like particles in rodents and rhesus macaques. *Intervirology* 1-2 (39): 93-103.
116. Yao, Q., *et al.*, **2002**. Intranasal immunization with SIV virus-like particles (VLPs) elicits systemic and mucosal immunity. *Vaccine* 19-20 (20): 2537-45.

117. Ciarlet, M., *et al.*, **1998**. Subunit rotavirus vaccine administered parenterally to rabbits induces active protective immunity. *J Virol* 11 (72): 9233-46.
118. Perez, N., *et al.*, **2006**. Rectal immunization with rotavirus virus-like particles induces systemic and mucosal humoral immune responses and protects mice against rotavirus infection. *J Virol* 4 (80): 1752-61.
119. Sedlik, C., *et al.*, **1997**. Recombinant parvovirus-like particles as an antigen carrier: a novel nonreplicative exogenous antigen to elicit protective antiviral cytotoxic T cells. *Proc Natl Acad Sci U S A* 14 (94): 7503-8.
120. Hu, Y.-C., *et al.*, **2003**. Generation of chimeric baculovirus with histidine-tags displayed on the envelope and its purification using immobilized metal affinity chromatography. *Enzyme and Microbial Technology* 4 (33): 445-452.
121. Slupetzky, K., *et al.*, **2007**. A papillomavirus-like particle (VLP) vaccine displaying HPV16 L2 epitopes induces cross-neutralizing antibodies to HPV11. *Vaccine* 11 (25): 2001-10.
122. Adams, M.R., Little, C.L., and Easter, M.C., **1991**. Modelling the effect of pH, acidulant and temperature on the growth rate of *Yersinia enterocolitica*. *J Appl Bacteriol* 1 (71): 65-71.
123. Teixeira, A.P., *et al.*, **2007**. Hybrid elementary flux analysis/nonparametric modeling: application for bioprocess control. *BMC Bioinformatics* (8): 30.
124. Teixeira, A.P., *et al.*, **2008**. Combining metabolic flux analysis tools and <sup>13</sup>C NMR to estimate intracellular fluxes of cultured astrocytes. *Neurochem Int* 3 (52): 478-86.
125. Licari, P. and Bailey, J.E., **1992**. Modeling the population dynamics of baculovirus-infected insect cells: Optimizing infection strategies for enhanced recombinant protein yields. *Biotechnol Bioeng* 4 (39): 432-441.
126. Roldão, A., *et al.*, **2006**. Intracellular dynamics in rotavirus-like particles production: Evaluation of multigene and monocistronic infection strategies. *Process Biochemistry* 10 (41): 2188-2199.
127. Asenjo, A.J., *et al.*, **2007**. A discrete mathematical model applied to genetic regulation and metabolic networks. *J Microbiol Biotechnol* 3 (17): 496-510.
128. Rasmussen, S., **2004**. Modelling of discrete spatial variation in epidemiology with SAS using GLIMMIX. *Comput Methods Programs Biomed* 1 (76): 83-9.
129. Roldão, A., *et al.*, **2007**. Modeling rotavirus-like particles production in a baculovirus expression vector system: Infection kinetics, baculovirus DNA replication, mRNA synthesis and protein production. *J Biotechnol* 4 (128): 875-894.
130. Dee, K.U. and Shuler, M.L., **1997**. Optimization of an assay for baculovirus titer and design of regimens for the synchronous infection of insect cells. *Biotechnol Prog* 1 (13): 14-24.
131. Power, J.F., *et al.*, **1996**. The kinetics of baculovirus adsorption to insect cells in suspension culture. *Cytotechnology* 2 (21): 155-163.
132. Gotoh, T., Chiba, K., and Kikuchi, K., **2004**. Probabilistic Characterization for Baculovirus-Infected Insect Cells Destined to Synthesize Progeny Viruses and Recombinant Protein and to Die. *J Chem Eng Jpn* 11 (37): 1357-1366.
133. Chen, L., *et al.*, **2000**. Hybrid modelling of biotechnological processes using neural networks. *Contr. Eng. Pract.* (8): 821-827.
134. Thompson, M.L. and Kramer, M.A., **1994**. Modeling chemical processes using prior knowledge and neural networks. *AIChE Journal* 8 (40): 1328-1340.
135. Psychogios, D.C. and Hungar, L.H., **1992**. A hybrid neural network – first principles approach to process modelling. *AIChE Journal* 10 (38): 1499-1511.
136. Jang, J.D., *et al.*, **2000**. Structured modeling of recombinant protein production in batch and fed-batch culture of baculovirus-infected insect cells. *Cytotechnology* 1-2 (34): 71-82.
137. Labbe, M., *et al.*, **1991**. Expression of rotavirus VP2 produces empty corelike particles. *J Virol* 6 (65): 2946-52.
138. Prasad, B.V., *et al.*, **1988**. Three-dimensional structure of rotavirus. *J Mol Biol* 2 (199): 269-75.
139. Crawford, S.E., *et al.*, **1999**. Heterotypic protection and induction of a broad heterotypic neutralization response by rotavirus-like particles. *J Virol* 6 (73): 4813-4822.
140. Jiang, B., Gentsch, J.R., and Glass, R.I., **2008**. Inactivated rotavirus vaccines: a priority for accelerated vaccine development. *Vaccine* 52 (26): 6754-8.
141. Volkman, L.E. and Goldsmith, P.A., **1985**. Mechanism of neutralization of budded Autographa Californica Nuclear Polyhedrosis Virus by a monoclonal antibody: Inhibition of entry by adsorptive endocytosis. *Virology* (143): 185-195.
142. Rapp, J.C., Wilson, J.A., and Miller, L.K., **1998**. Nineteen baculovirus open reading frames, including LEF-12, support late gene expression. *J Virol* 12 (72): 10197-206.
143. Guarino, L.A. and Summers, M.D., **1987**. Nucleotide Sequence and Temporal Expression of a Baculovirus Regulatory Gene. *J Virol* 7 (61): 2091-2099.



144. Passarelli, A.L. and Miller, L.K., **1993**. Identification of genes encoding late expression factors located between 56.0 and 65.4 map units of the *Autographa californica* nuclear polyhedrosis virus genome. *Virology* 2 (197): 704-14.
145. Guarino, L.A. and Summers, M.D., **1986**. Functional mapping of a trans-activating gene required for expression of a baculovirus delayed-early gene. *J Virol* 2 (57): 563-71.
146. Kool, M., *et al.*, **1994**. Identification of genes involved in DNA replication of the *Autographa californica* baculovirus. *Proc Natl Acad Sci U S A* 23 (91): 11212-6.
147. Lu, A. and Miller, L.K., **1995**. The roles of eighteen baculovirus late expression factor genes in transcription and DNA replication. *J Virol* 2 (69): 975-82.
148. Stewart, T.M., *et al.*, **2005**. The *Autographa californica* multiple nucleopolyhedrovirus ie0-ie1 gene complex is essential for wild-type virus replication, but either IE0 or IE1 can support virus growth. *J Virol* 8 (79): 4619-29.
149. Gong, M., Jin, J., and Guarino, L.A., **1998**. Mapping of ORF121, a factor that activates baculovirus early gene expression. *Virology* 2 (244): 495-503.
150. Wu, Y. and Carstens, E.B., **1998**. A baculovirus single-stranded DNA binding protein, LEF-3, mediates the nuclear localization of the putative helicase P143. *Virology* 1 (247): 32-40.
151. Passarelli, A.L. and Miller, L.K., **1993**. Identification and characterization of lef-1, a baculovirus gene involved in late and very late gene expression. *J Virol* 6 (67): 3481-8.
152. McDougal, V.V. and Guarino, L.A., **1999**. *Autographa californica* nuclear polyhedrosis virus DNA polymerase: measurements of processivity and strand displacement. *J Virol* 6 (73): 4908-18.
153. Okano, K., Mikhailov, V.S., and Maeda, S., **1999**. Colocalization of baculovirus IE-1 and two DNA-binding proteins, DBP and LEF-3, to viral replication factories. *J Virol* 1 (73): 110-9.
154. Lin, G., Slack, J.M., and Blissard, G.W., **2001**. Expression and localization of LEF-11 in *Autographa californica* nucleopolyhedrovirus-infected Sf9 cells. *J Gen Virol Pt 9* (82): 2289-94.
155. Morris, T.D., *et al.*, **1994**. Identification of lef-7: a baculovirus gene affecting late gene expression. *Virology* 2 (200): 360-9.
156. Knebel-Morsdorf, D., Kremer, A., and Jahnel, F., **1993**. Baculovirus gene ME53, which contains a putative zinc finger motif, is one of the major early-transcribed genes. *J Virol* 2 (67): 753-8.
157. Passarelli, A.L., Todd, J.W., and Miller, L.K., **1994**. A baculovirus gene involved in late gene expression predicts a large polypeptide with a conserved motif of RNA polymerases. *J Virol* 7 (68): 4673-8.
158. Guarino, L.A., *et al.*, **1998**. A virus-encoded RNA polymerase purified from baculovirus-infected cells. *J Virol* 10 (72): 7985-91.
159. Lin, G. and Blissard, G.W., **2002**. Analysis of an *Autographa californica* multicapsid nucleopolyhedrovirus lef-6-null virus: LEF-6 is not essential for viral replication but appears to accelerate late gene transcription. *J Virol* 11 (76): 5503-14.
160. McLachlin, J.R. and Miller, L.K., **1994**. Identification and characterization of vlf-1, a baculovirus gene involved in very late gene expression. *J Virol* 12 (68): 7746-56.
161. Ke, J., *et al.*, **2008**. *Autographa californica* multiple nucleopolyhedrovirus ac66 is required for the efficient egress of nucleocapsids from the nucleus, general synthesis of preoccluded virions and occlusion body formation. *Virology* 2 (374): 421-31.
162. Russell, R.L. and Rohrmann, G.F., **1997**. Characterization of P91, a protein associated with virions of an *Orgyia pseudotsugata* baculovirus. *Virology* 1 (233): 210-23.
163. Wilson, M.E. and Price, K.H., **1988**. Association of *Autographa californica* nuclear polyhedrosis virus (AcMNPV) with the nuclear matrix. *Virology* 1 (167): 233-41.
164. Blissard, G.W. and Rohrmann, G.F., **1989**. Location, sequence, transcriptional mapping, and temporal expression of the gp64 envelope glycoprotein gene of the *Orgyia pseudotsugata* multicapsid nuclear polyhedrosis virus. *Virology* 2 (170): 537-55.
165. Harrison, R.L., Jarvis, D.L., and Summers, M.D., **1996**. The role of the AcMNPV 25K gene, "FP25," in baculovirus polh and p10 expression. *Virology* 1 (226): 34-46.
166. Gross, C.H., *et al.*, **1993**. A baculovirus encoded 16-kDa glycoprotein localizes near the nuclear membrane of infected cells. *Virology* 1 (192): 386-90.
167. Wu, W., *et al.*, **2008**. *Autographa californica* multiple nucleopolyhedrovirus 38K is a novel nucleocapsid protein that interacts with VP1054, VP39, VP80, and itself. *J Virol* 24 (82): 12356-64.
168. Reilly, L.M. and Guarino, L.A., **1994**. The pk-1 gene of *Autographa californica* multinucleocapsid nuclear polyhedrosis virus encodes a protein kinase. *J Gen Virol* (75 ( Pt 11)): 2999-3006.
169. Kuzio, J., *et al.*, **1984**. Nucleotide sequence of the p10 polypeptide gene of *Autographa californica* nuclear polyhedrosis virus. *Virology* 2 (139): 414-8.
170. Roelvink, P.W., *et al.*, **1992**. Dissimilar expression of *Autographa californica* multiple nucleocapsid nuclear polyhedrosis virus polyhedrin and p10 genes. *J Gen Virol* (73 ( Pt 6)): 1481-9.
171. Kasman, L.M. and Volkman, L.E., **2000**. Filamentous actin is required for lepidopteran nucleopolyhedrovirus progeny production. *J Gen Virol Pt 7* (81): 1881-8.

172. De Gooijer, C.D., *et al.*, **1992**. A structured dynamic model for the baculovirus infection process in insect-cell reactor configurations. *Biotechnol Bioeng* 4 (40): 537-548.
173. Kumar, A. and Shuler, M., **1995**. Model of a split-flow air lift bioreactor for attachment-dependent, baculovirus-infected insect cells. *Biotechnol. Prog.* (11): 412-419.
174. Power, J., *et al.*, **1992**. Modelling the growth and protein production by insect cells following infection by a recombinant baculovirus in suspension culture. *Cytotechnology* 1-3 (9): 149-155.
175. Power, J.F., *et al.*, **1994**. Modeling and optimization of the baculovirus expression vector system in batch suspension culture. *Biotechnol Bioeng* 6 (44): 710-719.
176. Valentine, R.C. and Allison, A.C., **1959**. Virus particle adsorption. I. Theory of adsorption and experiments on the attachment of particles to non-biological surfaces. *Biochim Biophys Acta* (34): 10-23.
177. Petricevich, V.L., *et al.*, **2001**. Parameters that determine virus adsorption kinetics: toward the design of better infection strategies for the insect cell - baculovirus expression system. *Enzyme Microb Technol* 1 (29): 52-61.
178. Dee, K.U. and Shuler, M.L., **1997**. A mathematical model of the trafficking of acid-dependent enveloped viruses: Application to the binding, uptake, and nuclear accumulation of baculovirus. *Biotechnol Bioeng* 5 (54): 468-490.
179. Rosinski, M., Reid, S., and Nielsen, L.K., **2002**. Kinetics of baculovirus replication and release using real-time quantitative polymerase chain reaction. *Biotechnol Bioeng* 4 (77): 476-480.
180. Haas, R. and Nielsen, L.K., **2005**. A physiological product-release model for baculovirus infected insect cells. *Biotechnol Bioeng* 6 (91): 768-72.
181. Palomares, L.A., Lopez, S., and Ramirez, O.T., **2002**. Strategies for manipulating the relative concentration of recombinant rotavirus structural proteins during simultaneous production by insect cells. *Biotechnol Bioeng* 6 (78): 635-44.
182. Licari, P. and Bailey, J.E., **1991**. Factors influencing recombinant protein yields in an insect cell-baculovirus expression system: Multiplicity of infection and intracellular protein degradation. *Biotechnology and Bioengineering* 3 (37): 238-246.

# Chapter 2

## R<sub>e</sub>combinant baculovirus titration

This chapter was based on the following manuscript:

António Roldão, Rui Oliveira, Manuel J. T. Carrondo, Paula M. Alves (2009). "Error assessment in recombinant baculovirus titration: evaluation of different methods", *J Virol Methods*, 159(1): 69-80.

## Summary

The success of baculovirus/insect cells system in heterologous protein expression depends on the robustness and efficiency of the production workflow. It is essential that process parameters are controlled and include as little variability as possible. The multiplicity of infection (MOI) is the most critical factor since irreproducible MOIs caused by inaccurate estimation of viral titers hinder batch consistency and process optimization. This lack of accuracy is related to intrinsic characteristics of the method such as the inability to distinguish between infectious and non-infectious baculovirus. In this study, several methods for baculovirus titration were compared. The most critical issues identified were the incubation time and cell concentration at the time of infection. These variables influence strongly the accuracy of titers and must be defined for optimal performance of the titration method. Although the standard errors of the methods varied significantly (7% to 36%), titers were within the same order of magnitude; thus, viral titers can be considered independent of the method of titration. A cost analysis of the baculovirus titration methods used in this study showed that the alamarblue, real time Q-PCR and plaque assays were the most expensive techniques. The remaining methods cost on average 75% less than the former methods. Based on the cost, time and error analysis undertaken in this study, the end-point dilution assay, microculture tetrazolium assay and flow cytometric assay were found to be the techniques that combine better all these three main factors. Nevertheless, it is always recommended to confirm the accuracy of the titration either by comparison with a well characterized baculovirus reference stock or by titration using two different methods and verification of the variability of results.

## CONTENTS

<b>1. Introduction.....</b>	<b>54</b>
<b>2. Materials and methods .....</b>	<b>57</b>
2.1. Cell culture and baculovirus amplification .....	57
2.2. Baculovirus titration methods .....	58
2.3. Statistical analysis .....	65
<b>3. Results.....</b>	<b>66</b>
3.1. TCID <sub>50</sub> for baculovirus with and without GFP .....	66
3.2. MTT assay .....	69
3.3. Alamarblue assay .....	71
3.4. Cell size assay .....	72
3.5. Growth cessation assay .....	73
3.6. Real time Q-PCR.....	74
3.7. Plaque assay .....	74
3.8. Flow cytometry assay .....	75
3.9. Statistical analysis .....	76
<b>4. Discussion .....</b>	<b>78</b>
4.1. Critical issues in baculovirus titration .....	78
4.1.1. Effect of cell origin .....	78
4.1.2. Effect of CCI and baculovirus inoculation procedure .....	78
4.1.3. Effect of incubation time .....	80
4.1.4. Effect of defective interfering particles .....	81
4.1.5. Mathematical analysis .....	81
4.2. Error assessment in baculovirus titration methods .....	82
4.3. Estimated accuracy of titers .....	82
4.4. Methods cost, time and error analysis .....	83
<b>5. Acknowledgements.....</b>	<b>85</b>
<b>6. References .....</b>	<b>85</b>

## 1. INTRODUCTION

Heterologous protein expression implies the introduction of a recombinant gene into an expression host cellular system using appropriate vector systems allowing the maintenance and the expression of the gene of interest in the host cell. Within the different systems available today, the baculovirus/insect cells system has gained interest due to its high expression capacity and easy culture conditions. Nonetheless, its success depends on the robustness and efficiency of the production workflow. For this purpose, the critical process parameters must be controlled. The multiplicity of infection (MOI) is the most critical factor since irreproducible MOIs caused by inaccurate estimation of viral titers have a profound impact on the process performance and batch consistency. Indeed, underestimated titers result in higher MOIs than expected, thus leading to a severe metabolic burden effect which causes low protein productivity. Similarly, overestimated titers result in low production levels due to non-synchronous infection [1]. Thus, the selection of a titration method capable of distinguishing between infectious and non-infectious virus is essential to reduce errors of titers and MOI variability.

Traditional baculovirus titration methods are based on cell viability [2-4], plaque infectivity [5-7] and antibody-based assays [8; 9]. Within cell based methods, end-point dilution (TCID<sub>50</sub>), microculture tetrazolium (MTT), cell size and growth cessation assays are used most commonly. All these methods rely on cellular morphological changes caused by infection such as the decrease of cell viability and the increase of cell size. The extent of those phenomena is dependent on the virus dose and thus can be correlated with viral titers. The main advantage of plaque infectivity assay over cell based methods is the production of countable events, namely plaque forming units (pfu) *per* ml, which is the appropriate form to represent viral titers. However, the specificities of this method make difficult the

assessment of accurate viral titers. Antibody-based methods present even greater challenges since in many cases immunoreagents and antibodies against baculoviral proteins are not available commercially or are very expensive. All these methods, directly or indirectly, rely on host and cell-specific infectivity, yielding significant errors due to their semi-quantitative nature.

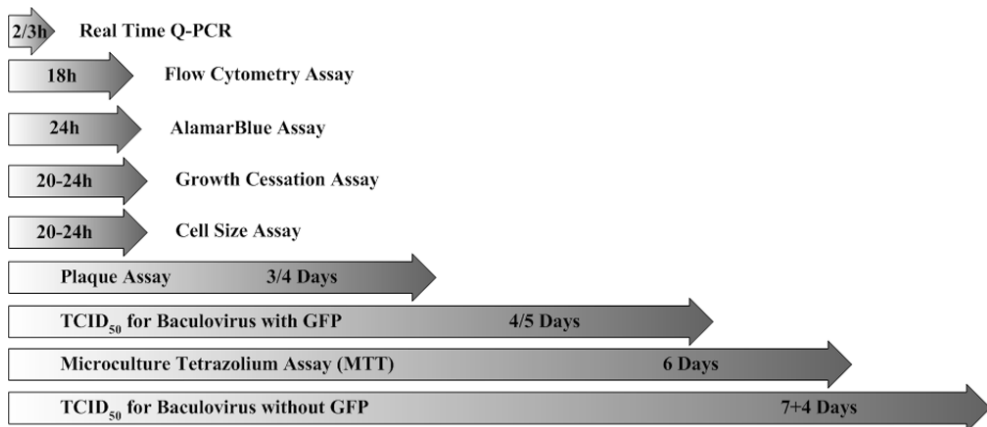
In an effort to increase accuracy, cell viability techniques using reporter proteins [10] and colorimetric indicators [11] have been developed. The use of fluorescent proteins such as the green fluorescent protein (GFP), which expression results from viral infection, improves the detection of infected cells (green fluorescence foci) and reduces both the viral titer variability and the titration time. Likewise, the use of dyes, such as the alamarBlue<sup>®</sup>, allows more rapid estimation of titers than standard cell viability assays since the correlation between cytopathic effect caused by infection and titer is assessed easily. Nonetheless, neither the presence of reporter proteins is always desirable nor the use of expensive dyes is always feasible. In addition, these techniques are labour intensive. Thus, novel baculovirus titration methods using real time Q-PCR [12-14] and flow cytometry [15-17] were developed recently. Real time Q-PCR assay is based on the intensity of a fluorescent signal generated during PCR amplification. By comparing the signal intensity obtained for standards of known viral concentrations with the signal intensity of the viral stock sample, viral titers are estimated. However, this requires the design of primers targeting a specific strand of nucleic acids corresponding to the protein of interest and the respective standards, which can be very complex. In addition, artefacts such as primer-dimers and defective interfering particles (non-infectious particles) decrease the accuracy of the titers measured. Flow cytometric assays involve the staining of viral genome or coat proteins of baculovirus with highly fluorescent specific dyes. The fluorescence intensity of stained virus stocks when compared

with the fluorescence intensity of calibration standards allows the estimation of viral titers. However, the lack of sensitivity, instability and insolubility in aqueous media of many dyes compromise the accuracy of titers. Some dyes are also difficult to synthesize and expensive. Recently, Kärkkäinen *et al.* (2009) developed a titration method able to correlate the percentage of cells expressing GFP to the baculovirus titers calculated by TCID<sub>50</sub>. The disadvantage of this technique is the fact that it involves the expression of a reporter protein which is not desirable for protein production processes designed for human use.

In summary, some titration methods show great potential to reduce the variability of viral titers but do not decrease the labour and titration times significantly. On the other hand, other methods reduce labour and titration time, but do not improve the accuracy of the method of titration. These inconsistencies stress that no method is ideal *per se* and the diversity of methods is still large. Indeed, depending on the virus/system specificities and available financial/technical resources, the proper titration method to use may differ. It is important that the selected method estimates viral titers with sufficient accuracy to control better the MOI. A comparative study to evaluate the accuracy of titers and cost/time-effectiveness would be valuable considering the enormous diversity of existing titration techniques.

In this study, recombinant *Autographa californica* multiple nucleopolyhedrovirus (AcMNPV) were amplified by infecting *Spodoptera frugiperda* Sf-9 cells at an MOI of 0.01 pfu/cell and titrated using different methods (Figure 7). Mathematical formulations and statistical tools were used to determine and compare the titers obtained by each method. Specificities and inter-relationships between various methods were studied to determine the method presenting the best association between time, cost and accuracy.





**Figure 7.** Overview of recombinant baculovirus titration methods and time required to perform each method.

## 2. MATERIALS AND METHODS

### 2.1. Cell culture and baculovirus amplification

*Sf-9* and *Sf-21* cell lines were obtained from American type culture collection (ATCC, Washington D.C., USA) and were cultivated in serum free media SF900II (Gibco, Glasgow, UK) at 27 °C in 500 ml erlenmeyers (50 ml working volume) (Duran, Mainz, Germany) stirred at 90 rpm and diluted routinely every four days using an inoculum of  $0.3 \times 10^6$  cell/ml. Recombinant AcMNPV coding for the main rotavirus proteins, vp2 (102.7 kDa [18]) fused to GFP, vp6 (44.9 kDa [19]) and vp7 (37.2 kDa [19]), were amplified by infecting *Sf-9* cells at a cell concentration of  $1 \times 10^6$  cell/ml and MOI of 0.1 pfu/cell in 250 ml working volume spinner flasks (Wheaton Science, New Jersey, USA) stirred at 150 rpm and at 27 °C. Baculoviruses were harvested at 60-70% of cell viability, approximately 96-120 hours post- infection (hpi), by centrifugation at  $200 \times g$  for 15 min at 4 °C and viral stocks were prepared. Stock A and B correspond to baculovirus coding for vp2-GFP, stock C for vp6 and stock D for vp7. The baculovirus were kindly

provided by Dr. Jean Cohen (Centre National pour la Recherche Scientifique-Institut National Recherche Agronomique, Gif-sur-Yvette, France). All genes were under the control of the polyhedrin promoter.

## **2.2. Baculovirus titration methods**

### **2.2.1. TCID<sub>50</sub> for baculovirus with and without GFP**

Different end-point dilutions were used. Stocks of baculovirus coding for rotavirus proteins fused to GFP (Stock A and B) were titrated according to Cha et al. (1997) whereas those without GFP (Stocks C and D) were titrated based on standard protocols [7; 20; 21].

In the TCID<sub>50</sub> method for baculovirus with GFP, the detection of infection is easier and more rapid than in standard methods since green fluorescence foci are indicative of infection. Briefly, 10 µl of viral stocks diluted 10<sup>-2</sup> to 10<sup>-8</sup> times were added to 100 µl of 1×10<sup>5</sup> cell/ml *Sf*-9 cells seeded previously onto a 96-well tissue culture plate (Sarstedt, Newton, USA). Plates were incubated at 27 °C for four to five days and checked for infection using a Leica DM IRB inverted microscope (470/480nm wavelength) (Leica Microsystems, Wetzlar, Germany).

TCID<sub>50</sub> for baculovirus without GFP was based on standard methods. Briefly, 100 µl of 2×10<sup>5</sup> cell/ml *Sf*-9 cells were seeded onto a 96-well tissue culture plate (Sarstedt, Newton, USA) and allowed to settle for 1 h at 27 °C. In each well, 10 µl of viral stocks diluted 10<sup>-2</sup> to 10<sup>-8</sup> times were added. Plates were incubated at 27 °C for seven days and cell morphology was observed using a Leica DM IRB inverted microscope (Leica Microsystems, Wetzlar, Germany). Positive/negative wells were determined based on the increase of the diameter of the cell and on inhibition of cell growth. However, since these morphological changes are difficult to observe by an “untrained eye”, a second round of infection was necessary. Therefore, 20

$\mu\text{l}$  of each viral suspension was added to 80  $\mu\text{l}$  of Sf-9 cells ( $2.5 \times 10^5$  cell/ml) seeded previously onto a 96-well tissue culture plate (Sarstedt, Newton, USA). Plates were incubated at 27 °C for four additional days and visualized using a Leica DM IRB inverted microscope (Leica Microsystems, Wetzlar, Germany) for detection of foci of infection.

Virus titers were calculated using the mathematical equations described by Reed and Muench (1938), Possee and King (1992) and Darling et al. (1998) (see Table 5).

### 2.2.2. MTT assay

The cell viability assay was performed as described by Mena et al. (2003). Briefly, 100  $\mu\text{l}$  of  $5 \times 10^5$  cell/ml of Sf-21 cells were seeded onto a 96-well tissue culture plate (Sarstedt, Newton, USA) and were allowed to settle for 1 h at 27 °C. Supernatant was removed and 100  $\mu\text{l}$  of viral stocks diluted  $10^{-1}$  to  $10^{-10}$  times was added *per* well and plates incubated six days at 27 °C. Thiazolyl blue tetrazolium bromide (Sigma-Aldrich, St. Louis, USA) was added (10% (v/v) of total volume *per* well at 5 mg/ml) and plates incubated at 27 °C for 4 h. The supernatant was removed and formazan crystals solubilised by adding 150  $\mu\text{l}$ /well of dimethyl sulfoxide (Sigma-Aldrich, St. Louis, USA). Plates were agitated for 10-20 min in a wellmix shaker WM-506 (Denley, Needham, USA) and absorbance (570/690nm wavelength) measured using a SPECTRAmax™ microplate reader (Molecular Devices Corporation, Sunnyvale, USA).

Collected data was analyzed using Prism 4 for Windows (GraphPad Software Inc., La Jolla, USA) to determine the tissue culture lethal dose 50 (TCLD<sub>50</sub>) (see Table 5). The conversion of TCLD<sub>50</sub> to viral titers (pfu/ml) was carried out using the mathematical regressions reported in this study (see section 2.2.1) and in Mena et al. (2003).

**Table 5.** Mathematical equations for titer estimation.

		$PD = \frac{[50\% - (\% \text{ rate of dilution next below } 50\%)]}{[(\% \text{ rate of dilution next above } 50\%) - (\% \text{ rate of dilution next below } 50\%)]}$ $\log(\text{endpoint}) = \log(\text{dilution next to and below } 50\%) + PD \times \log(\text{dilution factor})$ $pfu/ml = \frac{\text{endpoint}}{V} \times 0.69$	
	Reed and Muench (1938)		
TCID <sub>50</sub> (baculovirus w/o GFP and w/GFP)	Possee and King (1992)	$PD = \frac{[(\% \text{ rate of dilution next above } 50\%) - 50\%]}{[(\% \text{ rate of dilution next above } 50\%) - (\% \text{ rate of dilution next below } 50\%)]}$ $\log(TCID_{50}) = \log(\text{dilution next to and above } 50\% \text{ positives}) - PD$ $pfu/ml = \frac{TCID_{50}}{V} \times 0.69$	<p>PD – proportional distance</p> <p>V – volume of added virus</p> <p>X<sub>p</sub> – log of highest dilution inducing 100% infection</p> <p>d – log of dilution factor</p> <p>Σp<sub>i</sub> – sum of rates of positive results</p>
	Darling et al. (1998)	$\log(TCID_{50}) = X_p + \frac{d}{2} - d \sum p_i$ $pfu/ml = \frac{TCID_{50}}{V} \times 0.69$	
	MTT assay	$A = A_{\text{bottom}} + \frac{A_{\text{top}} - A_{\text{bottom}}}{1 + \left( \frac{dil}{EC_{50}} \right)^{\text{slope}}}$ $TCLD_{50} = \frac{1}{EC_{50} \times V}$ $\log(pfu/ml) = 0.829 \times \log(TCLD_{50}) + 2.512$	<p>A – absorbance</p> <p>A<sub>top</sub> – maximum absorbance (set to 100 after data normalization)</p> <p>A<sub>bottom</sub> – baseline absorbance (set to 0 after data normalization)</p> <p>EC<sub>50</sub> – dilution that provokes a response halfway between maximum and baseline</p> <p>dil – virus dilution</p>
	AlamarBlue assay	$I = \frac{R_{\text{virus},24} - R_{\text{cells},0}}{R_{\text{cells},24} - R_{\text{cells},0}} \times 100$ $I = I_{\text{bottom}} + \frac{I_{\text{top}} - I_{\text{bottom}}}{1 + \left( \frac{dil}{EC_{50}} \right)^{\text{slope}}}$ $TCLD_{50} = \frac{1}{EC_{50} \times V}$ $\log(pfu/ml) = 0.643 \times \log(TCLD_{50}) + 5.867$	<p>I – percentage of growth inhibition</p> <p>R<sub>virus,24</sub> – percentage reduction of alamarBlue® in infected cells for t=24hpi</p> <p>R<sub>cells,24</sub> – percentage reduction of alamarBlue® in viable cells for t=24hpi</p> <p>R<sub>cells,0</sub> – percentage reduction of alamarBlue® in viable cells for t=0hpi</p> <p>I<sub>top</sub> – maximum growth inhibition percentage</p> <p>I<sub>bottom</sub> – baseline growth inhibition percentage</p>

**Table 5 (cont.).** Mathematical equations for titer estimation.

Method of moments	$EP_{IC,n} = \frac{d_{C,n} - d_C}{d_{IC} - d_C}$	$EP_{IC,n}$ – estimation of infected cells proportion at the nth dilution $d_{C,n}$ – diameter of viable cells at the nth dilution $d_C$ – average diameter of viable cells $d_{IC}$ – average diameter of infected cells $\alpha, \beta$ – parameters $\in \mathbb{R}^2$ $[C]_{C,n}$ – concentration of viable cells at the nth dilution $[C]_{IC,n}$ – concentration of infected cells at the nth dilution
Maximum likelihood estimates	$d_{C,n} = (1 - EP_{IC,n}) \left[ \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(d_{C,n} - d_C)^2}{2\sigma^2}\right) \right] \times (EP_{IC,n}) \left[ \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(d_{C,n} - d_{IC})^2}{2\sigma^2}\right) \right]$	
Logistic function	$EP_{IC,n} = \frac{1}{1 + \exp(\beta[\ln(dil) - \alpha])}$ $[C]_{IC,n} = EP_{IC,n} \times [C]_{C,n}$ $pfu/ml = \frac{[C]_{IC,n}}{dil}$	
Cell size assay		
Growth cessation assay	$I = 100 \cdot \frac{[C]_{IC,24} - [C]_{C,10}}{[C]_{C,24} - [C]_{C,10}} \times 100$	$[C]_{IC,24}$ – concentration of infected cells for t=24hpi
	$I = I_{top} + \frac{I_{top} - I_{bottom}}{1 + \left(\frac{dil}{TCID_{50}}\right)^{slope}}$	$[C]_{C,40}$ – concentration of viable cells for t=0hpi
	$pfu/ml = \frac{[C]_{C,10}}{TCID_{50}} \times 0.69$	$[C]_{C,24}$ – concentration of viable cells for t=24hpi
Plaque assay	$pfu/ml = \frac{N_p \times V}{dil}$	$N_p$ – number of plates in a well at the nth dilution
Flow cytometry assay	$FC = \frac{(\% \text{ cells expressing GFP}) \times N_{c,well}}{V} \times dil$ $pfu/ml = FC \times 1.8$	$FC$ – titer determined by flow cytometry $N_{c,well}$ – number of cells per well

### 2.2.3. Alamarblue assay

In this study, an experimental design similar to that described by Pouliquen et al. (2006) was adopted. Briefly, 100  $\mu$ l of  $3 \times 10^4$  cell/ml Sf-21 cells seeded previously onto a 96-well tissue culture plate (Sarstedt, Newton, USA) were infected with 100  $\mu$ l of serial dilutions (1/2 to 1/1458 times) of virus stock. AlamarBlue<sup>®</sup> (Sigma-Aldrich, St. Louis, USA) was added to plates (10% (v/v)) and plates were incubated at 27 °C. The percentage of alamarBlue<sup>®</sup> reduction at 0 and 24 hpi was determined by measuring the fluorescence intensity and absorbance with an FL500 microplate fluorescence reader (530<sub>excitation</sub>/590<sub>emission</sub>nm wavelength) (Bio-Tek Instruments, Winooski, USA) and a SPECTRAmax<sup>™</sup> microplate reader (570/690nm wavelength) (Molecular Devices Corporation, Sunnyvale, USA), respectively, for  $t_0$  and  $t_{24}$  values assessment.

Data was analyzed by Prism 4 for Windows (GraphPad Software Inc., La Jolla, USA) and titers calculated using the mathematical equations described by Pouliquen et al. (2006) (see Table 5).

### 2.2.4. Cell size assay

Viral titers can be estimated based on the increase of the diameter of cell caused by infection as described by Janakiraman et al. (2006). Briefly, 10 ml of Sf-9 cells at  $1 \times 10^6$  cell/ml cultured in 125 ml erlenmeyers (20ml working volume) (Duran, Mainz, Germany) at 27 °C were infected with 10 ml of viral serial dilutions ranging from  $10^{-1}$  to  $10^{-4}$  times. Samples at time 0 and 24 hpi were collected and the viable cell diameter and the concentration were measured using a Casy<sup>®</sup>1 Cell counter plus Analyzer System Model TTC (Innovatis AG, Reutlingen, Germany).

Viral titers were determined using the solver function in Microsoft excel (Microsoft<sup>®</sup>, Redmond, USA) by applying different methods for parameter

estimation: method of moments, maximum likelihood estimates and the logistic function (see Table 5).

### **2.2.5. Growth cessation assay**

Cells normally stop growing immediately after infection at MOIs higher than 5 [22; 23]. Viral titers can be determined using this property since the inhibition of cell growth is dependent on the virus dose and thus related to the amount of virus infecting a given cell population. Briefly, 10 ml of Sf-9 cells at  $1 \times 10^6$  cell/ml cultured in a set of 125 ml erlenmeyers (20ml working volume) (Duran, Mainz, Germany) were infected with 10 ml of virus serial dilutions ranging from  $10^{-1}$  to  $10^{-4}$  times. Positive and negative controls corresponding to cells infected with undiluted virus stock and to uninfected cells, respectively, were also performed. Erlenmeyers were incubated for 20-24 h at 27 °C and cell concentration determined by counting cells on a fuchs-rosenthal haemocytometer (Brandt, Wertheim/Main, Germany) using the trypan blue exclusion dye (Merck, Darmstadt, Germany) method or by Casy<sup>®</sup>1 Cell counter plus Analyzer System Model TTC (Innovatis AG, Reutlingen, Germany).

Raw data was analyzed using the 4-parameter method, which assumes a sigmoidal inhibition curve with lower and upper asymptote and an estimated TCID<sub>50</sub>, implemented in microsoft excel (Microsoft<sup>®</sup>, Redmond, USA) and solved using the solver function. Viral titers were estimated by multiplying the ratio seeding density/TCID<sub>50</sub> by the conversion factor 0.69 [7] (see Table 5).

### **2.2.6. Real time Q-PCR**

Titers can be estimated within a couple of hours (2 to 3 h) using real time Q-PCR as described by Vieira et al. (2005). Briefly, viral DNA extraction of samples from stocks A to D was performed using the high pure viral nucleic

acid kit (Roche Diagnostics, Mannheim, Germany). Purified baculovirus DNA were analyzed by real-time quantitative PCR in LightCycler<sup>®</sup>1.5 instrument (Roche Applied Science, Indianapolis, USA) using Fast start DNA master SYBR Green I kit (Roche Diagnostics, Mannheim, Germany) and sequence-specific primers for the genes coding for the viral proteins vp2, vp6 and vp7 [13]. For each recombinant baculovirus stock, viral titers were estimated in pfu/ml using the standard calibration curve of cross points vs log concentrations of the purified DNA standard with known concentration (pfu/ml).

### 2.2.7. Plaque assay

The plaque assay technique was carried out as described by Possee and King (1992). Briefly, 2 ml of Sf-21 cells at a concentration of  $1 \times 10^6$  cell/ml were seeded onto a Nunclon<sup>™</sup> Surface 6-well plate (Nunc<sup>®</sup>, Langenselbold, Germany) and allowed to attach for 1 h. Meanwhile, serial viral dilutions were prepared ( $10^{-5}$  to  $10^{-8}$  times) in serum free media SF900II (Gibco, Glasgow, UK) supplemented with 10% of fetal calf serum (Gibco, Glasgow, UK). The supernatant was removed and 500 µl of virus dilutions were added to the plates. Plates were incubated at 27 °C for 1-4 h, the time period needed for virus adsorption and internalization [24] before removal of the supernatant. Two millilitres of an agarose overlay (1 part 3% (w/v) SeaPlaque<sup>®</sup> low gelling temperature agarose (Cambrex Bio Science Rockland, Inc., Rockland, USA) plus 2 parts serum free media SF900II (Gibco, Glasgow, UK) supplemented with 10% of fetal calf serum (Gibco, Glasgow, UK)) were added to each well. Plates were incubated for 3 to 4 days at 27 °C, before staining with 500 µl of Neutral Red (Sigma-Aldrich, St. Louis, USA). Any excess of Neutral Red was removed 24 h later and plates were further incubated at room temperature for plaque visualization.

Titers were estimated according to the method described by Possee and



King (1992) (see Table 5).

### 2.2.8. Flow cytometry assay

In this study, viral titers of stocks A and B were estimated by flow cytometry using a mathematical equation that correlates the percentage of cells expressing GFP with the concentration of baculovirus at infection [15]. Baculovirus stocks C and D do not express any reporter protein and were not considered for titration by flow cytometry. Briefly, 500  $\mu$ l of Sf-9 cells at  $1.5 \times 10^6$  cell/ml seeded onto 2.2 ml storage plates (ABgene Limited, Epsom, UK) were infected with 500  $\mu$ l of serial virus dilutions. It is important to stress that virus dilutions must be carefully selected since only those which induce a percentage of cells expressing GFP below 10% are considered for titer calculation. Thus, for titers estimated to be around  $10^8$  pfu/ml, dilutions from 1/300 to 1/9600 times are preferable while for those around  $10^7$  pfu/ml, the ideal dilutions are 1/150 to 1/4800 times. After incubation of plates at 27 °C for 18 h in an IKA KS 260 orbital shaker at 370 rpm (Sigma-Aldrich, St. Louis, USA), cell suspensions were transferred to 3.5 ml tubes (Sarstedt, Newton, USA) and the percentage of cells expressing GFP determined using a CyFlow® space flow cytometer (Partec, Görlitz, Germany).

Titers were estimated according to the method described by Kärkkäinen et al. (2009) (see Table 5).

## 2.3. Statistical analysis

For each viral stock (A-D) and titration method used in this study, replicates were performed, varying from a minimum of  $n = 3$  to a maximum of  $n = 11$ . All titers were expressed as the mean  $\pm$  standard deviation (SD). A statistical hypothesis test was used to evaluate if the titers estimated by each method were equal or not:  $H_0$  (Null hypothesis):  $\text{titer}_{\text{Method1}} = \text{titer}_{\text{Method2}}$

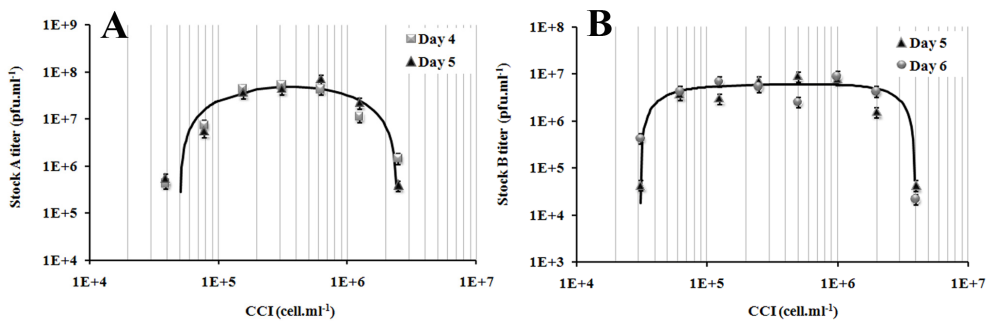
= ... = titer<sub>Method9</sub> versus  $H_1$  (alternative hypothesis): not all titers were equal. The correct statistical test to use considering the number of methods involved was the one-way ANOVA. This analysis was based on the assumption that errors were independently, identically, and normally distributed. We could confirm this by performing two tests: 1) the Shapiro-Wilk test to confirm that the distribution of values in each method followed a normal distribution – the normality is verified for p-value ( $p > \alpha = 0.05$ ); 2) the Levene's test to verify that the variance of data in each method was the same – variances are considered equal for  $p > \alpha = 0.05$ . After verified this assumption, one-way ANOVA was performed in SPSS 13.0 for Windows (SPSS Inc., Chicago, USA) for a level of confidence of 95% ( $\alpha = 0.05$ ). According to the resulting p-value, the null hypothesis was rejected ( $p < \alpha = 0.05$ ) or not ( $p > \alpha = 0.05$ ); the higher the p-value the more evidence we had that titers were not statistically different. In case of rejection of  $H_0$ , the method estimating titers statistically different from the mean was identified by performing a multiple comparison test – Sheffé test. This test consisted in comparing the titer obtained by a certain method with the remaining ones; for example, the titer of method 1 (M1) with M2, M1 with M3, M1 with M4 and thereafter. For each set of comparisons, the correspondent p-values specified if the null hypothesis was to be rejected or not ( $H_0$ : titer<sub>M1</sub> = titer<sub>M2</sub> versus  $H_1$ : the two titers were not equal). The null hypothesis was rejected for  $p < \alpha = 0.05$ . The method presenting p-values  $< \alpha = 0.05$  in all comparisons was that estimating titers statistically different from the mean.

### 3. RESULTS

#### 3.1. TCID<sub>50</sub> for baculovirus with and without GFP

Based on the literature, there are two major variables that have to be considered when using TCID<sub>50</sub> method [10]: cell concentration at the time

of infection (CCI) and virus-cell incubation time. In order to determine the effect of CCI on estimated titers, cells at a seeding density between  $3 \times 10^4$  and  $4 \times 10^6$  cell/ml were infected with recombinant baculovirus of stocks A and B and titers determined as described in section 2.2.1. Our results showed that for CCIs between  $1 \times 10^5$  and  $1 \times 10^6$  cell/ml there was no significant difference in titers obtained for both stocks (Figure 8). To test the effect of the incubation time on estimated titers, independent studies were performed for baculovirus with and without GFP since the two titration protocols are based on different assumptions for identification of positive/negative wells.



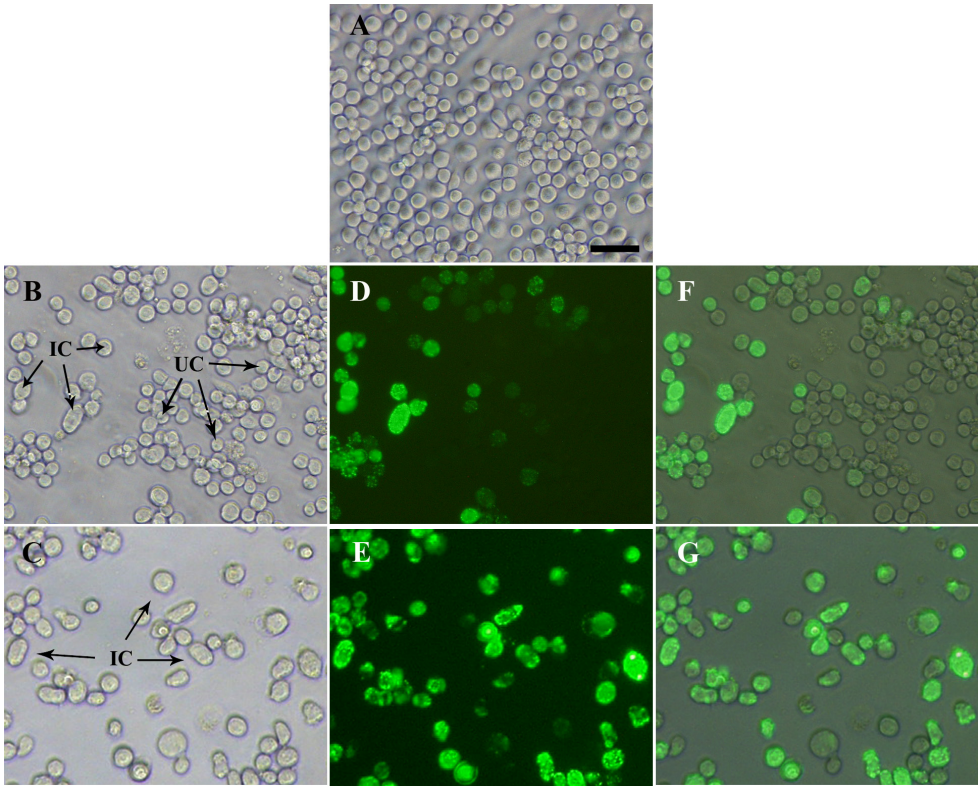
**Figure 8.** TCID<sub>50</sub> method for baculovirus with GFP: effect of cell concentration at the time of infection (CCI) and incubation time on viral titers of stock A (A) and stock B (B). Caption denotes the incubation time, error bars denote standard errors and the lines represent curve fittings.

In the first study for baculovirus with GFP, different incubation times were tested (2-8 days) using two strategies: 1. performing the infection without removing the supernatant of the seeded cells; 2. discarding the supernatant before virus addition. In strategy 1, titers varied around 18% while in strategy 2 titers variability reached 68%. For the incubation time, the titers variability of stocks A and B was on average 17% and 4%, respectively, for 4, 5 and 6 days of incubation; for the remaining times tested, the variability of stocks A and B titers increased to 57% and 39% respectively (data not

shown). Thus, the effect of incubation time on the accuracy of titers can be considered negligible for the time-frame of 4 to 6 days of incubation. Based on these results, further experiments using the recombinant baculovirus with GFP (Stock A and B) were performed at a CCI of  $1 \times 10^5$  cell/ml, using strategy 1 and 4 to 5 days of incubation.

In the second study for baculovirus without GFP, four incubation times were tested (4, 5, 6 and 7 days). Results indicate that the detection of positive/negative wells was only possible from day seven as the inhibition of cell growth and the increase of cell size became visible when compared to uninfected cells used as negative control (Figure 9A vs 9B). Nevertheless, this screening was difficult and highly susceptible to errors since the cell morphological changes were not markedly perceptible even for a “trained eye”. Thus, the infection of a second 96-well plate with the virus solution from the plate one was required to increase the accuracy of titers. Consequently, it was necessary to evaluate the effect of different incubation times (4, 5, 6 and 7 days) on final titers in this second infection procedure. The results indicated that four days of incubation was sufficient to identify clearly infection foci. Indeed, we observed a more pronounced inhibition of cell growth and a notorious increase of cell size combined to the loss of the characteristic round cellular shape in cells infected for 7+4 days compared to both uninfected and seven days infected cells (Figures 9A-C). These findings were confirmed using a baculovirus with GFP as indicated in figures 9D and 9E. Therefore, further experiments using the recombinant baculovirus without GFP were performed at a CCI of  $2 \times 10^5$  cell/ml using method 1 and incubation time of 7+4 days.

The mathematical equations available in literature for viral titer estimation (see Table 5) were compared. Results showed that the variability in titer yields was on average 9-17%, demonstrating that any of the mathematical equations could be used without compromising final titer estimations.



**Figure 9.** Assessment of positive/negative wells using the TCID<sub>50</sub> method: screening of uninfected (A) and infected cells with recombinant baculovirus from stock A (B to G). Figures B and C show infection foci identification based on the inhibition of cell growth and changes of shape and size of infected vs uninfected cells, for days 7 and 7+4 respectively. The presence of green fluorescence due to GFP allows an easier identification of the infected cells (Figures D and E). Figures F and G are the respective overlapping of Figures B with D and C with E. IC, infected cells; UC, uninfected cells. Scale bar 50  $\mu$ m.

Data collected from a run of experiments were analysed and viral stock titers assessed (Table 6). Titer yields were within the same log,  $10^7$  pfu/ml, and standard errors between 18 and 35%.

### 3.2. MTT assay

Two major points were evaluated in this method: the effect of CCI on TCLD<sub>50</sub> values and the correlation between TCLD<sub>50</sub> and pfu/ml.

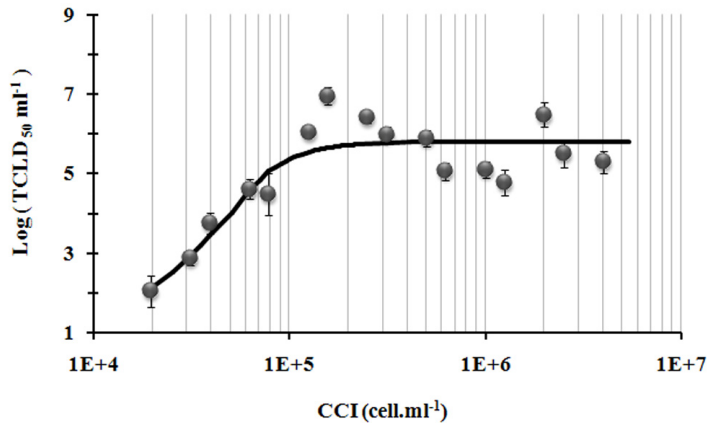
**Table 6.** Comparison of different methods for baculovirus titration.

	Baculovirus stock titers (pfu.ml <sup>-1</sup> ) × 10 <sup>7</sup>				Method standard error (average)
	Stock A	Stock B	Stock C	Stock D	
TCID <sub>50</sub> for baculovirus w/o GFP	-	-	2.2 ± 0.8	3.5 ± 0.7	27%
TCID <sub>50</sub> for baculovirus w/GFP	1.7 ± 0.3	1.1 ± 0.3	-	-	22%
MTT assay	1.5 ± 0.3	2.5 ± 0.8	1.2 ± 0.1	0.9 ± 0.2	20%
Alamarblue assay	0.9 ± 0.3	1.1 ± 0.3	5 ± 1	1.4 ± 0.4	28%
Cell size assay	2.3 ± 0.4	1.0 ± 0.1	6.2 ± 0.5	14 ± 3	13%
Growth cessation assay	1.4 ± 0.5	50 ± 18	4.4 ± 1.6	2.4 ± 0.9	36%
Real time Q-PCR	741 ± 134	279 ± 6	55 ± 7	451 ± 34	10%
Plaque assay	1.63 ± 0.06	0.54 ± 0.05	0.30 ± 0.02	0.57 ± 0.06	7%
Flow cytometry assay	2.7 ± 0.3	1.6 ± 0.2	-	-	13%
<b>Stock standard error (average)</b>	20%	28%	27%	29%	

To define the CCI range within which no significant differences in TCLD<sub>50</sub> values were observed, *Sf*-21 cells seeded at densities ranging from 2×10<sup>4</sup> to 4×10<sup>6</sup> cell/ml were infected with recombinant baculovirus of stock A. Results showed two completely different patterns: a linear relationship between TCLD<sub>50</sub> and CCI at low cell concentration (10<sup>4</sup> to 10<sup>5</sup> cell/ml) and an apparent TCLD<sub>50</sub> plateau at CCIs higher than 10<sup>5</sup> cell/ml (Figure 10). Thus, further experiments were performed at a CCI of 5×10<sup>5</sup> cell/ml.

To allow direct comparison with other methods, it was necessary to correlate the TCLD<sub>50</sub> values obtained by MTT assay with pfu/ml. Therefore, parallel MTT and TCID<sub>50</sub> assays were performed at CCIs from 2×10<sup>4</sup> to 4×10<sup>6</sup> cell/ml. By representing log (pfu/ml) vs log (TCLD<sub>50</sub>/ml), we determined a linear regression with a slope of 0.8 ± 0.2 and a y-intercept of 3 ± 1 ( $r^2 = 0.85$ ) that allowed the estimation of viral stock titers (Table 6). The titers obtained were within the same log, 10<sup>7</sup> pfu/ml, and methods intra-variability between 12 and 30% which is in agreement with literature [4; 25]. Interestingly, an average difference of 27% was observed between

these titers and those calculated using the equation proposed by Mena et al. (2003). This emphasises the need to assess the correlation between  $TCLD_{50}$  and pfu/ml for accurate titer estimation.



**Figure 10.** Microculture tetrazolium (MTT) assay for baculovirus from stock A: effect of cell concentration at the time of infection (CCI) on  $TCLD_{50}/ml$  values. Error bars denote standard errors and the lines represent curve fitting.

### 3.3. Alamarblue assay

In this technique, three major variables affected strongly viral titer estimation: cell concentration at the time of infection, cell line and incubation time.

An initial screening to assess the seeding densities inducing a linear alamarBlue<sup>®</sup> reduction response, indicative of optimal conditions, was performed. For this purpose, different concentrations ranging from  $1 \times 10^3$  to  $5 \times 10^5$  cell/ml were tested. The results showed an optimal linear relation for cell concentrations between  $1.5 \times 10^4$  cell/ml and  $3 \times 10^4$  cell/ml. To evaluate the precision of estimated titers at these two concentrations, *Sf*-9 and *Sf*-21 cells were infected with recombinant baculovirus from stock C and titers determined according to Pouliquen et al. (2006). Two incubation times, 24 and 34 h, were also tested. Results showed that titers obtained at  $1.5 \times 10^4$

cell/ml were on average 40% higher than those at  $3 \times 10^4$  cell/ml; however, their variability was much higher (45% vs 20%). Assays performed with *Sf*-9 cells were incoherent and the titers impossible to calculate, whereas the titers obtained using *Sf*-21 cells were consistent and presented significant precision. The reason underneath this discrepancy of results is not understood. For incubation times, titers obtained at 34 hpi were on average 65% lower than those at 24 hpi.

Based on these results, recombinant baculovirus of stocks A to D were titrated using *Sf*-21 cells at a CCI of  $3 \times 10^4$  cell/ml and incubation time of 24 h (Table 6). Titters were within the  $10^7$  pfu/ml yield and method variability between 26-32% which was in close agreement with the results obtained for other methods.

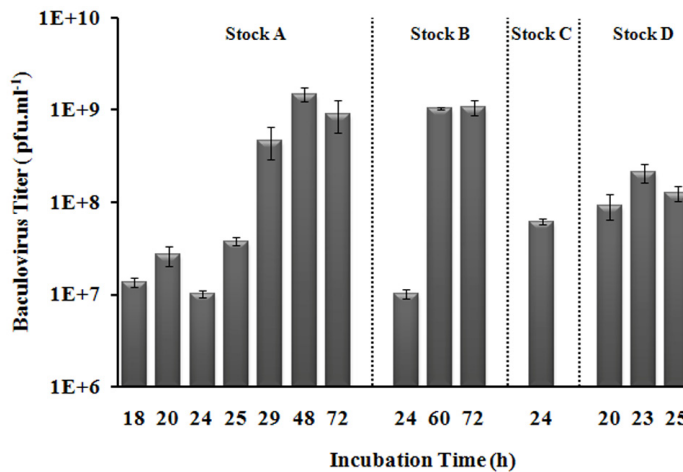
### **3.4. Cell size assay**

The definition of optimal incubation time was critical for this technique. Thus, *Sf*-9 cells infected with recombinant baculovirus from stocks A to D at a CCI of  $1 \times 10^6$  cell/ml with virus serial dilutions ranging from  $10^{-1}$  to  $10^{-4}$  times were incubated during different times (18-72 h) and titers were calculated as described in section 2.2.4. Results showed 2-log difference between titers obtained for incubation times around 18-24 h and higher than 24 h (see stocks A and B in Figure 11). These differences were related to the extent of the cell swelling phenomenon after infection caused by the different amount of virus existing at each incubation time tested. For incubation times higher than 24 h, virus underwent at least two replication cycles resulting in a higher amount of virus than the initial viral amounts. Consequently, the increase of cell size was more pronounced than expected leading to an overestimation of viral titers. On the other hand, for incubation times lower than 24 h, since virus were only able to complete one replication cycle, the cell swelling effect was mainly due to initial virus



dosage ensuring that the estimated titers reflect the real population of virus used for the infection. Based on these results, further experiments were performed with incubation times around 20-24 h.

The mathematical equations available in literature to determine the parameter  $EP_{IC,n}$  and consequently the titer of baculovirus stocks (see Table 5) were compared; an average variability around 16% was calculated. Overall, estimated titers were within  $10^7$  to  $10^8$  pfu/ml with standard errors below 17% (Table 6).



**Figure 11.** Cell size assay for baculovirus from stock A to D: effect of incubation time on recombinant baculovirus titers. Caption denotes the viral stock and error bars methods variability.

### 3.5. Growth cessation assay

Similarly to cell size assay, the most critical issue in this technique was the incubation time. Thus, to evaluate the effect of this variable on final titers, Sf-9 cells infected with recombinant baculovirus from stocks A to D at a CCI of  $1 \times 10^6$  cell/ml with virus serial dilutions ranging from  $10^{-1}$  to  $10^{-4}$  times were incubated during different times (24, 36, 48, 72 and 105 h) and titers determined as described in section 2.2.5. Results showed that titer yields varied significantly for the incubation times of 24 h and higher than 24 h (1

to 1.5-log difference). This observation was explained by the different extent of the cell growth inhibition phenomenon after infection caused by the different amounts of virus existing at each incubation time tested. Taking into account that baculovirus replication times range from 16 to 18 h, for incubation times lower than 24 h, virus were only able to complete one replication cycle. In these conditions, the inhibition of cell growth observed was attributed to the amount of virus initially used for the infection and thus directly correlated with the viral titer. Conversely, for incubation times higher than 24 h, the cell growth inhibition effect was more pronounced than expected since it was a consequence of the viral amounts used for the infection combined with the amount of newly replicated virus; consequently, in these conditions, the viral titers were overestimated. Based on these results, further experiments were performed using incubation times of 20 to 24 h.

Viral titers of all four recombinant baculovirus stocks were within the same order of magnitude,  $10^7$  pfu/ml with the exception of stock B (Table 6). The standard errors for this method were significant, with an average of 36%.

### **3.6. Real time Q-PCR**

Based on the method described in section 2.2.6, viral titers of baculovirus stocks A to D were determined (Table 6). Estimated titers were within  $10^8$ - $10^9$  pfu/ml and method standard errors between 2-18%. Interestingly, the titers obtained by real time Q-PCR were 1 to 2-log higher than those determined by the other techniques.

### **3.7. Plaque assay**

In this method, there were three major variables that influenced strongly titer estimation: incubation time, medium to use in the agarose overlay and cell concentration at the time of infection (CCI).

In order to define optimal conditions, several experiments were set up. Results showed that 3 to 4 days of incubation induced plaque formation with a reasonable size and the number of plaques being between 10 and 150 as recommended in Possee and King (1992). In addition, the use of serum free medium (SF900II) improved the quality and the definition of the plaques compared to the tissue culture medium (TC100). In terms of CCI, concentrations below  $10^6$  cell/ml yielded only few or no plaques (confluence was not achieved and cell detachment was visible). On the other hand, too high cell concentrations hindered plaque visualization because the cells reached confluence too soon and superposition between plaques and cells became too pronounced. Based on these results, further experiments were performed with a CCI of  $10^6$  cell/ml using serum free medium (SF900II) in the agarose overlay for 3 to 4 days of incubation.

Baculovirus stock titers determined by this method were lower than the titers determined by other techniques with the exception of the baculovirus stock A whose titer was within the same order of magnitude,  $10^7$  pfu/ml (Table 6). These lower titers may be related to non efficient virus/cell contacts or agarose overlay temperature, factors that affect strongly the amount of countable plaque forming units.

### **3.8. Flow cytometry assay**

This technique allowed titer estimation within one working day, similar to the growth cessation and cell size assays. However, titer precision was higher. Indeed, the error associated to flow cytometry assay, around 13%, was lower than the other two techniques, in the order of 17 to 35% (Table 6). In terms of stock titers, they were within the same order of magnitude,  $10^7$  pfu/ml. The ratio between infectious particles obtained by flow cytometry and TCID<sub>50</sub> was calculated to be 1.5, which is in agreement with what is reported in Shen et al. (2002).

### 3.9. Statistical analysis

In the case of stock A, both normality and variance equality were verified ( $0.05 < p < 0.69$ ). However, the one-way ANOVA results suggested that viral titers were statistically different ( $p = 2.77 \times 10^{-15} < \alpha = 0.05$ ), thus it was necessary to perform the Sheffé test to identify possible outliers. This multiple comparison test identified real time Q-PCR as the technique presenting significant differences in viral titers ( $p < 0.05$  for all multiple comparisons). A new statistical test performed excluding this technique suggested that there was no statistical difference between titers ( $p = 0.082 > \alpha = 0.05$ ). This means that viral titers determined by real time Q-PCR were statistically different from the average.

For stock B, one-way ANOVA was performed after verified the normality of data and the variance equality ( $0.14 < p < 0.48$ ). The resulting p-value,  $2.04 \times 10^{-43}$ , was lower than  $\alpha = 0.05$  meaning that viral titers were statistically different. The Sheffé test indicated real time Q-PCR technique as the method contributing most for such differences ( $p < 0.05$  for all multiple comparisons). Excluding this technique from further statistical analysis, the one-way ANOVA presented a p-value ( $p = 1.1 \times 10^{-21}$ ) above 0.05 suggesting that, apart from real time Q-PCR, another method was responsible for such significant variability. By performing a new Sheffé test, growth cessation assay was detected as that reporting viral titers significantly different from the rest ( $p < 0.05$  for all multiple comparisons) and consequently excluded from further statistical analysis. The new one-way ANOVA was performed and results indicated that there was sufficient evidence ( $p = 0.158 > \alpha = 0.05$ ) suggesting that there was no statistical difference between titers. This means that viral titers determined by real time Q-PCR and growth cessation assay were statistically different from the average.

Similarly, in stock C, two Sheffé tests were necessary for identification of outliers, namely real time Q-PCR and plaque assay. Excluding these two techniques from further statistical analysis, one-way ANOVA performed after verification of data normality and variance equality ( $0.06 < p < 0.66$ ) indicated that there was sufficient evidence ( $p = 0.065 > \alpha = 0.05$ ) to suggest that viral titers were statistically not significantly different. This means that viral titers determined by real time Q-PCR and plaque assay were statistically different from the average.

For stock D, two methods were excluded from further statistical analysis since both yielded p-values lower than 0.05 in the Sheffé test, namely the real time Q-PCR and the cell size assay. Multiple comparisons of the remaining techniques were assessed using the Sheffé test and all p-values were higher than 0.05 with the exception of the relation between TCID<sub>50</sub> and plaque assay, where  $p = 0.02 < \alpha = 0.05$ . Since this difference in viral titers was statistically significant, plaque assay method was not considered in further tests. The final one-way ANOVA test indicated that there was sufficient evidence ( $p = 0.055 > \alpha = 0.05$ ) to suggest that viral titers were not significantly different. This means that viral titers determined by real time Q-PCR, cell size assay and plaque assay were statistically different from the average.

In summary, this analysis indicates that the real time Q-PCR and the plaque assay were the techniques estimating viral titers statistically different from the average. The remaining techniques estimated similar titers for all baculovirus stocks tested. Therefore, the probability of estimating a viral titer statistically different from the real one using one of these two methods is significantly higher than with any other method tested in this study.

## 4. DISCUSSION

### 4.1. Critical issues in baculovirus titration

Baculovirus titration is a very complex task since results are highly dependent on the method used for titration. Often, the same viral stock titrated in different laboratories lead to different results; this lack of harmonization originates difficulties in the interpretation and comparison of results. Our aim was to identify the critical parameters affecting virus titer estimation for different analytical methods and procedures commonly used in laboratories worldwide.

#### 4.1.1. Effect of cell origin

The best cell line to use for baculovirus titration has to be capable of estimating viral titers with significant accuracy. *Sf*-21 and *Sf*-9 cells are usually used for titration but a clear utilization of one or the other cell line is not obvious since both cell lines are optimal for some titration methods and sub-optimal for others. Indeed, our results showed that some methods estimated titers with higher accuracy using *Sf*-9 cell line and other methods using *Sf*-21 (see Table 7). The reasons for these differences are not fully understood but cell size (*Sf*-21 cells have a wider size distribution than *Sf*-9) (O'Reilly et al. 1994), resistance to shear stress (*Sf*-9 cells are slightly less fragile than *Sf*-21), ability to form regular monolayers (*Sf*-21 cells form more irregular monolayers than *Sf*-9) [26] and different susceptibility to infection by baculovirus [27] might be factors responsible for these discrepant behaviours.

#### 4.1.2. Effect of CCI and baculovirus inoculation procedure

Optimal CCIs were dependent on the method of titration (see Table 7). This fact can be explained by the different detection methods used for identification of infection, either based on the expression of a foreign

**Table 7.** Summary of the evaluated methods for baculovirus titration.

Assay	Method type	Cell line	CCI (cell.ml <sup>-1</sup> )	Inoculum volume (V <sub>cells</sub> + V <sub>virus</sub> )	Incubation time	Detection method	Mathematical equations	References
<b>TCID<sub>50</sub></b>	Indirect	Sf-9	1×10 <sup>5</sup> (w/gfp) and 2×10 <sup>5</sup> (w/o gfp)	100 µl + 10 µl	4-5 days (w/gfp) and 7+4 days (w/o gfp)	Expression of GFP (w/gfp) and cell size, shape and growth inhibition (w/o gfp)	TCID <sub>50</sub> to pfu.ml <sup>-1</sup> (all mathematical equations estimate similar titers)	Cha et al. (1997) (w/gfp) and Possee and King (1992) (w/o gfp)
<b>MTT</b>	Indirect	Sf-21	5×10 <sup>5</sup>	100 µl + 100 µl	6 days	Cleavage of thiazolyl blue tetrazolium bromide by a mitochondrial enzyme in viable cells	TCID <sub>50</sub> to pfu.ml <sup>-1</sup>	Mena et al. (2003)
<b>Alamarblue</b>	Indirect	Sf-21	3×10 <sup>4</sup>	100 µl + 100 µl	24 h	Reduction of alamarBlue® by a mitochondrial enzyme in viable cells	TCID <sub>50</sub> to pfu.ml <sup>-1</sup>	Pouliquen et al. (2006)
<b>Cell size</b>	Indirect	Sf-9	1×10 <sup>6</sup>	10 ml + 10 ml	20-24 h	Increase of cell size caused by infection	EP <sub>IC<sub>50</sub></sub> to pfu.ml <sup>-1</sup> (all mathematical equations estimate similar titers)	Janakiraman et al. (2006)
<b>Growth cessation</b>	Indirect	Sf-9	1×10 <sup>6</sup>	10 ml + 10 ml	20-24 h	Cell growth inhibition caused by infection	TCID <sub>50</sub> to pfu.ml <sup>-1</sup>	-
<b>Real time Q-PCR</b>	Direct	-	-	2 µl (purified vDNA)	-	Viral particle identification by detection of viral DNA	pfu.ml <sup>-1</sup> (titers directly in pfu.ml <sup>-1</sup> due to the use of standard calibration curves)	Vieira et al. (2005)
<b>Plaque assay</b>	Direct	Sf-21	1×10 <sup>6</sup>	2 ml + 500 µl	3-4 days	Formation of plaque-forming units caused by infection	pfu.ml <sup>-1</sup> (titers directly in pfu.ml <sup>-1</sup> due to the formation of plaque-forming units)	Possee and King (1992)
<b>Flow cytometry</b>	Indirect	Sf-9	1.5×10 <sup>6</sup>	500 µl + 500 µl	18 h	Percentage of cells expressing GFP	Percentage of cells expressing GFP to pfu.ml <sup>-1</sup>	Karkkainen et al. (2009)

protein (TCID<sub>50</sub> w/GFP and flow cytometry assays), or cell viability (MTT and alamarblue assays), or cytopathic effects caused by infection (TCID<sub>50</sub> w/o GFP, cell size and growth cessation assays), or even plaque formation (plaque assay). The type of culture system is another possible explanation. Since some methods use static cultures whereas others require cells in suspension, the cell concentrations to be used for each method are completely different.

In relation to the inoculation procedure, results for TCID<sub>50</sub> demonstrated that the removal of the supernatant prior to the addition of viral serial dilutions had a negative impact on virus titers. Dehydration of attached cells, insufficient incubation times, plate type, medium exchange and cell concentration are possible reasons for this negative effect.

The volume of cell and baculovirus inoculum was different for the methods tested. Higher volumes for assays using suspension cultures (flow cytometry, cell size and growth cessation assays) and lower ones for static cultures (MTT, TCID<sub>50</sub>, plaque assay and alamarblue assays).

#### **4.1.3. Effect of incubation time**

Optimal incubation times varied from hours to days (Table 7). These differences are associated with the method used for the detection of the infection. The methods based on the expression of GFP (TCID<sub>50</sub> assay), the cleavage of thiazolyl blue tetrazolium bromide by a mitochondrial enzyme in viable cells (MTT assay) or the formation of plaque forming units (plaque assay) require high incubation times for identification of infection. Other techniques, based on cytopathic effects resulting from infection of cells by baculovirus (cell size and growth cessation assays) require lower incubation times. It is important to underline that, in these later cases, the incubation time is dependent on the replication cycle time of the virus. Incubation times higher than virus replication time lead to an overestimation



of viral titers since the inhibition of cell growth and the increase of cell size is the result of both initial and newly replicated virus. On the other hand, incubation times lower than the virus replication time restrains the estimation of viral titers since the viruses are unable to affect efficiently the cell growth or the increase of cell diameter.

#### **4.1.4. Effect of defective interfering particles**

Defecting interfering particles (DIPs), although lacking envelope or capsid proteins responsible for virus infectivity, contain genetic material. Techniques quantifying viral particles by detection of the viral DNA are unable to discriminate between infectious particles and DIPs and consequently will overestimate the viral titer. This was the case for the real time Q-PCR method presented here, that yielded titers that were significantly higher than the titers estimated by the other assays.

The propagation of baculovirus from a reference stock containing DIPs and/or accumulation of DIPs derived from incorrect propagation process increases exponentially the risk of generating a viral stock with significant percentage of these particles. Strategies can be implemented to avoid the accumulation of DIPs such as the addition of a purification step prior to viral titration. An alternative is to propagate the virus and quantify newly synthesized genomes, as described by [28]. Nonetheless, these strategies may impact negatively on final titers due to baculovirus aggregation associated with the purification step [29] or insufficient incubation time for the formation of progeny virus.

#### **4.1.5. Mathematical analysis**

Most of the methods presented in this study relied on specific mathematical correlations for the estimation of viral titers (indirect methods). The exceptions were the plaque assay and the real time Q-PCR which

quantified directly viral particles (direct methods) (see Table 7). To allow direct comparison between methods, it was necessary to standardize titers to pfu/ml. Therefore, correlations between viral genomes/ml and pfu/ml, in direct methods, and between TCID<sub>50</sub>, TCID<sub>50</sub>, % of cells expressing GFP and EP<sub>IC,n</sub> values and pfu/ml, in indirect methods, were used.

#### **4.2. Error assessment in baculovirus titration methods**

Based on the statistical analysis performed in section 3.9, the standard error for each stock was calculated. The results showed average standard errors between 20 to 29% which was acceptable considering how different the methods are. The standard error for each method was also estimated. Plaque assay and real time Q-PCR presented the lowest average standard errors (7% and 10% respectively) while alamarblue and growth cessation assays presented the highest ones (28% and 36% respectively). These results are in agreement with literature values and within acceptable viral titer variations defined by regulatory agencies ( $\pm 0.5 \log_{10}$ ) [4; 25; 30].

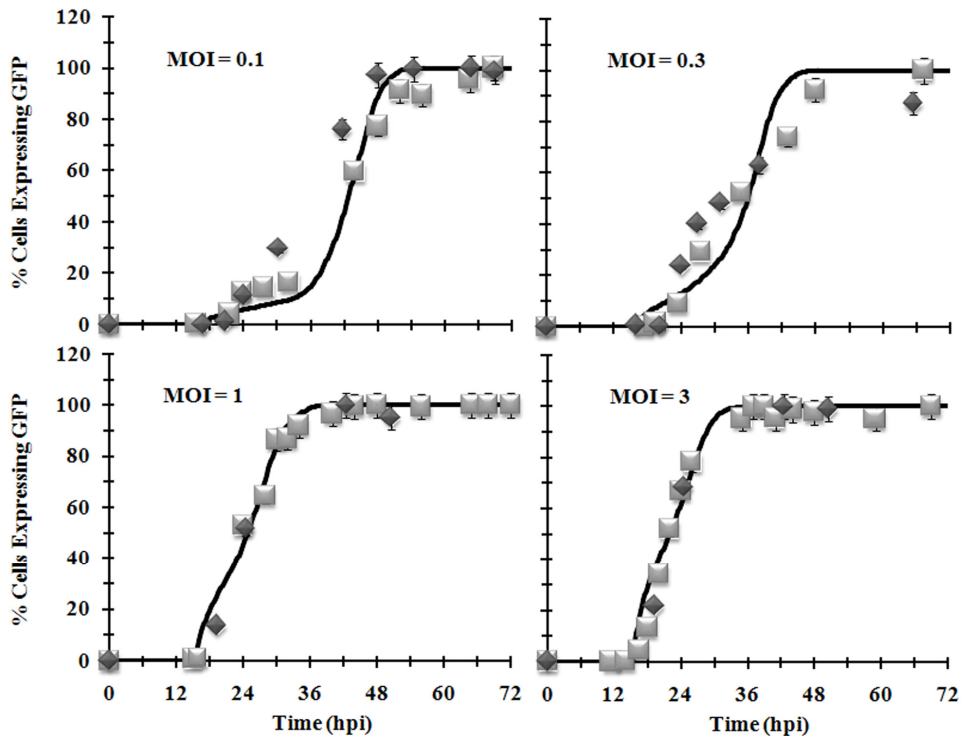
Despite the variability in the standard errors of the methods, viral titers were within the same order of magnitude,  $10^7$  pfu/ml. The exception was real time Q-PCR by the reasons mentioned in section 4.1.4. Thus, viral titers can be considered independent of the method of titration.

#### **4.3. Estimated accuracy of titers**

In order to verify the accuracy of estimated titers, two different runs of experiments (I and II) were performed. In each one, the infection kinetics of Sf-9 cells, at a cell concentration around  $10^6$  cell/ml, with stock A baculovirus at MOIs of 0.1, 0.3, 1 and 3 were followed. Collected samples were analysed by flow cytometry for assessment of percentage of cells expressing GFP.

Considering the error associated to process parameters CCI and MOI, both

profiles in experiments I and II can be considered identical (Figure 12). This means that the confidence interval,  $[-SD+titer, titer+SD]$ , for the stock A baculovirus is narrow, otherwise infection kinetics profiles would vary significantly. In addition, results prove that viral stock titers were estimated accurately.

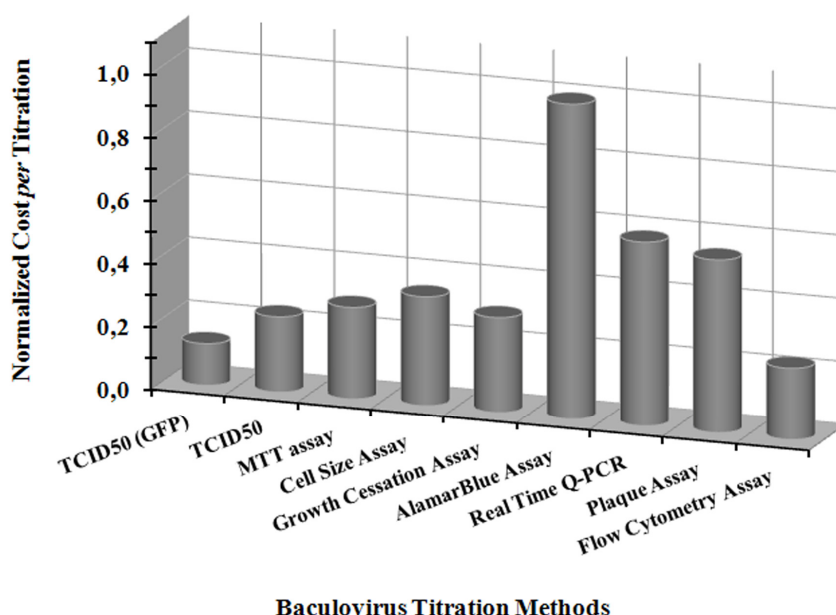


**Figure 12.** Estimated accuracy of titers: infection kinetics of Sf-9 cells, at a cell concentration of approximately 106 cell/ml, with stock A baculovirus at MOIs of 0.1, 0.3, 1 and 3. Caption denotes the MOI used and lines curve fittings. Symbols represent experimental data: experiment I (dark diamonds) and experiment II (grey squares).

#### 4.4. Methods cost, time and error analysis

A cost analysis of the baculovirus titration methods used in this study was performed (Figure 13). This analysis was based on the cost of consumables such as reagents, plastic ware, medium and other material necessary for the assays. The costs associated with equipment and human resources were not considered since these differ from lab to lab. The

alamarblue assay was the most expensive technique and TCID<sub>50</sub> method for baculovirus with GFP the least expensive one. The real time Q-PCR and plaque assay presented costs which were almost half of that of alamarblue assay. The costs associated with the remaining techniques were on average 75% lower than that of alamarblue assay.



**Figure 13.** Cost analysis of baculovirus titration methods. Titration costs were normalized for easier comparison. TCID<sub>50</sub> (GFP) and TCID<sub>50</sub> represent the end-point dilution method for baculovirus with and without GFP respectively.

Based on the cost, time and error analysis here performed, the techniques that combine better all these three main variables are the TCID<sub>50</sub>, MTT and the flow cytometric assays. Nevertheless, it is always recommended to confirm the accuracy of the titration either by comparing it with a well characterized baculovirus reference stock or by titration using two different methods and verification of the variability of results. Since standards stocks are not always easy to obtain, especially when dealing with a significant number of different baculovirus, the second alternative would allow for a better knowledge on the accuracy of estimated titers.

## 5. ACKNOWLEDGEMENTS

The authors wish to thank Dr. Vicente Bernal and Eng. Tiago Vicente for excellent advice and thoughtful discussions. This study was supported by the Portuguese Fundação para a Ciência e Tecnologia (POCTI/BIO/55975/2004 and SFRH/BD/21910/2005) and European Commission (Baculogenes, LSHB-2006-037541).

## 6. REFERENCES

1. Roldão, A., Carrondo, M.J.T., Alves, P.M., Oliveira, R., **2008**. Stochastic simulation of protein expression in the baculovirus/insect cells system. *Comp & Chem Eng.* 32 (1-2): 68-77.
2. Janakiraman, V., Forrest, W.F., Chow, B., Seshagiri, S., **2006**. A rapid method for estimation of baculovirus titer based on viable cell size. *J Virol Methods.* 132 (1-2): 48-58.
3. Lynn, D.E., **1992**. Improved efficiency in determining the titer of the *Autographa californica* baculovirus nonoccluded virus. *Biotechniques.* 13 (2): 282-285.
4. Mena, J., Ramirez, O.T., Palomares, L., **2003**. Titration of Non-Occluded Baculovirus Using a Cell Viability Assay. *Biotechniques.* 34 (2): 260-264.
5. Hink, W.F., Vail, P.V., **1973**. A plaque assay for titration of alfalfa looper nuclear polyhedrosis virus in a cabbage looper (TN-368) cell line. *J. Invertebr. Pathol.* 22 168-174.
6. O'Reilly, D.R., Miller, L.K., Luckow, V.A., **1994**. Baculovirus expression vectors: A laboratory manual, first ed. Oxford University Press, Inc., New York.
7. Possee, R.D., King, L.A., **1992**. The Baculovirus Expression System: A laboratory guide, first ed. Chapman & Hall, London.
8. Kitts, P.A., Green, G., **1999**. An immunological assay for determination of baculovirus titers in 48 hours. *Anal Biochem.* 268 (2): 173-178.
9. Kwon, M.S., Dojima, T., Toriyama, M., Park, E.Y., **2002**. Development of an antibody-based assay for determination of baculovirus titers in 10 hours. *Biotechnol Prog.* 18 (3): 647-651.
10. Cha, H.J., Gotoh, T., Bentley, W.E., **1997**. Simplification of titer determination for recombinant baculovirus by green fluorescent protein marker. *Biotechniques.* 23 (5): 782-786.
11. Poulighen, Y., Kolbinger, F., Geisse, S., Mahnke, M., **2006**. Automated baculovirus titration assay based on viable cell growth monitoring using a colorimetric indicator. *Biotechniques.* 40 (3): 282-292.
12. Hitchman, R., Siaterli, E., Nixon, C., King, L., **2006**. Quantitative real-time PCR for rapid and accurate titration of recombinant baculovirus particles. *Biotechnol Bioeng.* 96 (4): 810-814.
13. Vieira, H.L., Estevao, C., Roldao, A., Peixoto, C.C., Sousa, M.F., Cruz, P.E., Carrondo, M.J., Alves, P.M., **2005**. Triple layered rotavirus VLP production: kinetics of vector replication, mRNA stability and recombinant protein production. *J Biotechnol.* 120 (1): 72-82.
14. Lo, H.R., Chao, Y.C., **2004**. Rapid titer determination of baculovirus by quantitative real-time polymerase chain reaction. *Biotechnol Prog.* 20 (1): 354-360.
15. Kärkkäinen, H.-R., Lesch, H.P., Määttä, A.I., Toivanen, P.I., Mähönen, A.J., Roschier, M.M., Airenne, K.J., Laitinen, O.H., Ylä-Herttuala, S., **2009**. A 96-well format for a high-throughput baculovirus generation, fast titrating and recombinant protein production. *BMC Research Notes.* 2 (63): 1-6.
16. Stoffel, C.L., Kathy, R.F., Rowlen, K.L., **2005**. Design and characterization of a compact dual channel virus counter. *Cytometry A.* 65 (2): 140-147.
17. Shen, C.F., Meghrou, J., Kamen, A., **2002**. Quantitation of baculovirus particles by flow cytometry. *J Virol Methods.* 105 (2): 321-330.
18. Labbe, M., Charpilienne, A., Crawford, S.E., Estes, M.K., Cohen, J., **1991**. Expression of rotavirus VP2 produces empty corelike particles. *J Virol.* 65 (6): 2946-2952.

19. Prasad, B.V., Wang, G.J., Clerx, J.P., Chiu, W., **1988**. Three-dimensional structure of rotavirus. *J Mol Biol.* 199 (2): 269-275.
20. Darling, A.J., Boose, J.A., Spaltro, J., **1998**. Virus assay methods: accuracy and validation. *Biologicals.* 26 (2): 105-110.
21. Reed, L.J., Muench, H., **1938**. A simple method of estimating fifty per cent endpoints. *Am J Hyg.* 27 493-497.
22. Hu, Y.C., Bentley, W.E., **2001**. Effect of MOI ratio on the composition and yield of chimeric infectious bursal disease virus-like particles by baculovirus co-infection: deterministic predictions and experimental results. *Biotechniques.* 75 (1): 104-119.
23. Roldão, A., Vieira, H.L., Charpilienne, A., Poncet, D., Roy, P., Carrondo, M.J., Alves, P.M., Oliveira, R., **2007**. Modeling rotavirus-like particles production in a baculovirus expression vector system: Infection kinetics, baculovirus DNA replication, mRNA synthesis and protein production. *J Biotechnol.* 128 (4): 875-894.
24. Dee, K.U., Shuler, M.L., **1997**. Optimization of an assay for baculovirus titer and design of regimens for the synchronous infection of insect cells. *Biotechnol Prog.* 13 (1): 14-24.
25. LaBarre, D.D., Lowy, R.J., **2001**. Improvements in methods for calculating virus titer estimates from TCID<sub>50</sub> and plaque assays. *J Virol Methods.* 96 (2): 107-126.
26. Hink, W.F., Thomsen, D.R., Davidson, D.J., Meyer, A.L., Castellino, F.J., **1991**. Expression of three recombinant proteins using baculovirus vectors in 23 insect cell lines. *Biotechnol Prog.* 7 (1): 9-14.
27. Volkman, L.E., Summers, M.D., **1975**. Nuclear polyhedrosis virus detection: relative capabilities of clones developed from *Trichoplusia ni* ovarian cell line TN-368 to serve as indicator cells in a plaque assay. *J Virol.* 16 (6): 1630-7.
28. Ranheim, T., Mathis, P.K., Joelsson, D.B., Smith, M.E., Campbell, K.M., Lucas, G., Barmat, S., Melissen, E., Benz, R., Lewis, J.A., Chen, J., Schofield, T., Sitrin, R.D., Hennessey, J.P., Jr., **2006**. Development and application of a quantitative RT-PCR potency assay for a pentavalent rotavirus vaccine (RotaTeq). *J Virol Methods.* 131 (2): 193-201.
29. Michalsky, R., Pfromm, P.H., Czermak, P., Sorensen, C.M., Passarelli, A.L., **2008**. Effects of temperature and shear force on infectivity of the baculovirus *Autographa californica* M nucleopolyhedrovirus. *J Virol Methods.* 153 (2): 90-96.
30. FDA, **1998**. Guidance for Industry Q5A Viral Safety Evaluation of Biotechnology Products Derived From Cell Lines of Human or Animal Origin. 27.

# Chapter 3

## Stochastic infection and protein expression in the *AcMNPV/Sf-9* system

This chapter was based on the following manuscript:

António Roldão, Manuel J.T. Carrondo, Paula M. Alves and R. Oliveira (2008). "Stochastic simulation of protein expression in the baculovirus/insect cell system", *Computers and Chemical Engineering*, 32(1-2): 68-77.

## Summary

In this paper, a stochastic dynamical model for heterologous protein expression in the baculovirus/insect cell system is presented. The model describes foreign protein expression under the control of the polyhedrin promoter, which is the most commonly used promoter in this system. The present study explores the hypothesis of gene size being the main factor affecting the rate of protein expression in the host cell. The infection process, prior to protein expression, is of stochastic nature. Thus, the combination of infection and intracellular dynamics results in a complex stochastic/structured dynamical model which is very challenging in a systems engineering perspective. Due to the randomness of virus binding to the host cells, the intracellular dynamics of individual cells show unique patterns, potentially leading to a very large scale stochastic problem. The size of the problem was reduced by considering a finite number of subpopulations of “equal” cells. A discrete time formulation was adopted for realization of the stochastic infection and numerical integration. The method is illustrated by the application to the production of rotavirus protein  $vp_2$  in suspended cultures of *Spodoptera frugiperda* Sf-9 cells infected with recombinant *Autographa californica* multiple nucleopolyhedrovirus enclosing the gene coding for  $vp_2$ .



## CONTENTS

<b>1. Introduction.....</b>	<b>90</b>
<b>2. Proposed optimisation method .....</b>	<b>92</b>
2.1. Stochastic infection model.....	93
2.2. A structured model for intracellular protein expression.....	96
<b>3. Case study: production of vp2.....</b>	<b>98</b>
3.1. Model calibration .....	98
3.2. Stochastic simulation.....	100
3.3. Population dynamics .....	102
3.4. Intracellular dynamics.....	103
3.5. Productivity assessment.....	105
<b>4. Conclusions and future perspectives .....</b>	<b>107</b>
<b>5. Acknowledgements.....</b>	<b>108</b>
<b>6. References .....</b>	<b>108</b>

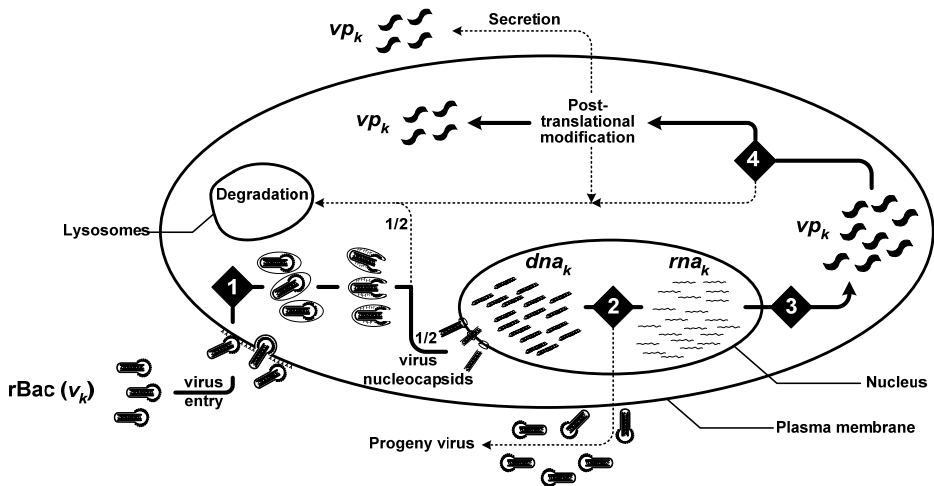
## 1. INTRODUCTION

The *Baculoviridae* are a family of occluded DNA viruses, which are pathogenic to some arthropods, namely to insects of the order *Lepidopteran*, and totally safe to humans. These viruses have a large, double-stranded, circular DNA genome of 88-200 kbp [1]. The use of baculovirus remotes to the early 1950s, when they were used as natural control agents of insect pest populations. At present, genetically engineered recombinant baculovirus are seen as powerful vectors for expressing heterologous proteins in insect cells, both *in vivo* and *in vitro*.

In heterologous protein production, one of the most used processes consists on the infection of Sf-9 cells with *Autographa californica* multiple nucleopolyhedrovirus (AcMNPV) [2-8]. This system offers some advantages over prokaryotic and other eukaryotic systems such as: 1. the very strong polyhedrin (polh) promoter with protein expression levels up to 500 mg.L<sup>-1</sup> [9]; 2. ability to express large proteins (> 50 kDa); 3. similar efficacy in protein folding and posttranslational modifications to mammalian cells; 4. low cost and 5. ability to express large cDNA inserts as well as multiple genes.

The main steps involved in the production of foreign proteins using the AcMNPV/Sf-9 system are depicted in Figure 14. Firstly, recombinant baculoviruses (rBac) containing the gene(s) of interest **k**, **v<sub>k</sub>**, bind to the host cell membrane mediated by cell surface receptors. About 1/2 of the viruses reach the host cell nucleus, with the remaining viruses being degraded in lysosomes. Details of AcMNPV trafficking in insect cells can be found elsewhere [10-13]. Once inside the host cell nucleus, viral genes are expressed in a well orchestrated cascade that is normally divided in three stages: early, late and very late gene expression [14]. Viral DNA (vDNA) replication starts in the late phase, around 6 hours post-infection (hpi), after

essential genes for vDNA replication are expressed in the early phase (1-6 hpi). The expression of foreign genes under the control of the polh promoter starts in the very late phase around 24 hpi. Also in the very late phase, some of the replicated vDNA, **dna<sub>k</sub>**, is encapsidated resulting in new virions, which then bud from the host cell. Virus budding is known to occur around 16 hpi and endures for almost 36 hours (h) [10; 15; 16]; the newly budded virus are able to start a second round of infection. The free vDNA is routed to the transcription pathway. The foreign genes are then transcribed into the corresponding mRNAs, **rna<sub>k</sub>** (step 2). In step 3, mRNA is translated into the corresponding polypeptide chains in the ribosomes, **vp<sub>k</sub>**. These newly synthesized proteins can be degraded or enter the secretion pathway (step 4). Once completed the post-translational modifications, the correctly folded proteins are released to the extracellular matrix or remain intracellularly.



**Figure 14.** Network of heterologous protein production in AcMNPV/Sf-9 system. Step 1: extracellular baculovirus, **v<sub>k</sub>**, enter the cell by adsorptive endocytosis. As the pH inside endosome drops, viruses' nucleocapsids containing the viral DNA are released to cytoplasm and migrate to cell nucleus. Once inside the nucleus, viral DNA replication is initiated, **dna<sub>k</sub>**. Step 2: viral DNA is routed to the transcription pathway where foreign genes are transcribed into the corresponding mRNAs, **rna<sub>k</sub>**. Step 3: mRNA leaves the nucleus to ribosomes for protein synthesis, **vp<sub>k</sub>**. Step 4: the secretory pathway delivers these newly synthesized proteins to the extracellular matrix.

The multiplicity of infection (MOI – number of virus *per* cell) and time of infection (TOI) are key parameters for AcMNPV/Sf-9 system optimisation. It has been shown for several proteins that the maximum productivity is obtained at relatively low MOIs, within the range of 0.01 to 0.1 virus.cell<sup>-1</sup> [15-18]. At such MOIs, the statistics of infection becomes an important factor for process modelling. Licari and Bailey (1992) proposed an implicit stochastic model for monolayer Sf-9 cultures where the probability of infection was calculated by the Poisson distribution. Protein expression was defined empirically with zero order kinetics with the time window for protein expression dependent on the number of virions infecting the host cell. Their model confirmed that protein productivity has a maximum at low MOIs. A similar approach was followed by Hu and Bentley (2000, 2001) for suspended Sf-9 cultures. Protein expression was considered to follow zero order kinetics and the kinetic constant was defined as a function of the number of virions infecting the host cell. Roldão *et al.* (2006, 2007) proposed a structured model describing the intracellular dynamics of vDNA, mRNA and foreign protein expression. The model assumed a deterministic infection, which is acceptable for synchronous infection at high MOIs.

The main objective of the present work is to bind intracellular dynamics based on traditional mass-action kinetics with the stochastic infection process into a stochastic dynamical model valid for both low and high MOIs.

## 2. PROPOSED OPTIMISATION METHOD

In this section, a method is presented for the optimisation of the MOI in the AcMNPV/Sf-9 system. The method is confined to the expression of foreign proteins under the control of the polh promoter. When dealing with the

synthesis of large and complex particles, biological details become vital for process optimisation. Therefore, the more complex scenario of stochastic structured modelling was adopted. The main guidelines for this optimisation model are the following:

- Stochastic infection. Low MOIs are important for process optimisation; thus, the stochastic nature of the infection process is considered in this method;
- Structured protein expression. The intracellular dynamics of vDNA, mRNA and viral protein are considered;
- Gene size. The main distinguishing factor in the intrinsic protein expression kinetics of different proteins is the size of the underlying gene;
- Maximum translation capacity. This hypothesis states that there is a maximum translation capacity of mRNA blocks coming from the polyhedrin promoter.

## 2.1. Stochastic infection model

Due to the high computation power requirements, the Sf-9 cell population, with an initial cell *density*  $\mathbf{n(0)}$  (cell.ml<sup>-1</sup>), is partitioned in **npop** subpopulations with a cell count of  $\mathbf{n_i = n(0) / npop}$  (cell.ml<sup>-1</sup>) each. Then, the dynamics of cell density in subpopulation **i** is given by the equation:

$$\frac{dn_i}{dt} = \mu_i n_i - k_{d,i} n_i \quad \text{Eq. 3.1}$$

with  $\mu_i$  (h<sup>-1</sup>) the specific growth rate of subpopulation **i** defined as:

$$\mu_i = \begin{cases} \mu_{max}, & \text{if population 'i' is healthy} \\ 0, & \text{if population 'i' is infected} \end{cases} \quad \text{Eq. 3.2}$$

and  $k_{d,i}$  (h<sup>-1</sup>) the cell death rate of subpopulation **i** defined as:

$$k_{d,i} = \begin{cases} k_{d,int} + k_{d,inf}, & t > TOI_i + \delta^* \\ k_{d,int}, & otherwise \end{cases} \quad \text{Eq. 3.3}$$

In equation 3.3,  $k_{d,int}$  ( $h^{-1}$ ) is the intrinsic cell death rate of Sf-9 cells,  $k_{d,inf}$  ( $h^{-1}$ ) represents the increase in cell death rate due to infection,  $TOI_i$  (h) is the time of infection of subpopulation  $i$  and  $\delta^*$  (hpi) is the time required for the infection to accelerate host cell death. The increase in cell death rate due to infection is correlated with the intracellular viral load [19]:

$$k_{d,inf} = \Delta k_d \text{LOG}_{10} \left( \sum_k dna_k \right) \quad \text{Eq. 3.4}$$

with  $\Delta k_d$  ( $h^{-1}$ ) the increase in cell death rate due to infection correspondent to 10 intracellular vDNA copies and  $dna_k$  ( $dna.cell^{-1}$ ) the intracellular vDNA concentration delivered by virus  $k$  ( $k$  denotes the foreign gene in the virus).

The rate of virus  $k$  entry into subpopulation  $i$ ,  $r_{inf,k,i}$  ( $dna.cell^{-1}h^{-1}$ ) is defined as:

$$r_{inf,k,i} dt = \begin{cases} X_{k,i}, & t \leq TOI_i + \delta_{reinf} \\ 0, & t > TOI_i + \delta_{reinf} \end{cases} \quad \text{Eq. 3.5}$$

with  $X_{k,i}$  a random variable denoting the number of viruses  $k$  per unit cell ( $dna.cell^{-1}$ ) which penetrate the cells of subpopulation  $i$  in the time interval  $[t, t+dt]$  and  $\delta_{reinf}$  (hpi) the time period over which cells continue to be re-infected. It is here assumed that the random variable  $X_{k,i}$  follows the Poisson probabilistic model:

$$P_{X_{k,i}=nk} = \frac{\exp(-\lambda(t))\lambda(t)^{nk}}{nk!} \quad \text{Eq. 3.6}$$

In Licari and Bailey (1992), the shape parameter  $\lambda(t)$  of the Poisson distribution was defined as the “dynamic” MOI (the ratio of extracellular

virus to host cells at time  $t$ ) multiplied by an infection efficiency parameter ( $\alpha$ ). Here we follow a different approach. The kinetics of virus entry into Sf-9 cells can be described accurately by a first order adsorption equation [11; 20]. The stochastic infection characterized by equations 3.5 and 3.6 must be in accordance with this observation, confirmed experimentally by many authors. We assume that the first order attachment rate defines the average behaviour for the stochastic process. Thus the shape parameter is defined here as follows:

$$\lambda(t) = k_a v_k(t) dt \quad \text{Eq. 3.7}$$

with  $k_a$  ( $\text{ml} \cdot \text{cell}^{-1} \cdot \text{h}^{-1}$ ) the first order attachment constant and  $v_k$  ( $\text{dna} \cdot \text{ml}^{-1}$ ) the concentration of extracellular virus  $k$ . Equation 3.7 can be interpreted as the average number of virus binding to a single cell in the time interval  $[t, t+dt]$ . The dynamics of extracellular virus  $v_k$  is then given by:

$$\frac{dv_k}{dt} = \sum_{i=1}^{npop} (-r_{inf,k,i} + r_{bud,k,i}) n_i \quad \text{Eq. 3.8}$$

Extracellular virus concentration has two contributions: the depletion of virus due to infection and also virus budding, which are the newly synthesized virus in the host cells secreted to the extracellular medium. This latter term is discussed in the next section.

It should be noted that equations 3.1 to 3.8 take into account the phenomena of re-infection in a very straightforward way. In the work of Hu and Bentley (2000, 2001), the probability of re-infection is also calculated and used to correct the depletion of extracellular virus. But, because no explicit stochastic formulation was adopted, viruses coming from re-infection were assumed to be degraded in lysosomes. With the method adopted here, viruses arriving to the host cell nucleus through infection or re-infection are treated in the same way.

## 2.2. A structured model for intracellular protein expression

The equations presented in the lines below refer to a given subpopulation  $i$ . To simplify the notation, the index  $i$  is omitted hereon. The intracellular dynamics of vDNA delivered by virus  $k$ ,  $\mathbf{dna}_k$  ( $\text{dna.cell}^{-1}$ ), is defined by:

$$\frac{d \mathbf{dna}_k}{dt} = \eta_t r_{inf,k} + r_{dnarep,k} - r_{bud,k}, \quad \mathbf{dna}_k(0) = 0 \quad \text{Eq. 3.9}$$

The first term in the right-hand side is the rate of infection (see equation 3.5) multiplied by the trafficking efficiency,  $\eta_t$ . According to Dee *et al.* (1997a), half of the internalized virus are degraded in lysosomes, thus  $\eta_t = 0.5$ . The second term is the vDNA replication rate,  $r_{dnarep,k}$  ( $\text{dna.cell}^{-1}.\text{h}^{-1}$ ), in the host cell nucleus, here defined as:

$$r_{dnarep,k} = \begin{cases} k_{RDNA} \mathbf{dna}_k \left( 1 - \frac{\delta^* - (t - TOI)}{\delta^* - \delta_{DNA,low}} \right), & \delta_{DNA,low} \leq t \leq \delta^* \\ 0, & otherwise \end{cases} \quad \text{Eq. 3.10}$$

Three main features of vDNA replication process can be outlined: 1. vDNA replication follows first order kinetics, i.e., template limited replication where  $k_{RDNA}$  ( $\text{h}^{-1}$ ) represents the first-order vDNA replication constant; 2. there is a well defined time window  $(t - TOI) \in [\delta_{DNA,low}, \delta^*]$  for vDNA replication, where  $\delta_{DNA,low}$  (hpi) represents the time for its beginning and 3. the rate of vDNA replication decreases linearly in time. This linear decrease is attributed to the metabolic burden caused by the synthesis of virus related molecules [15; 16; 18; 19]; vDNA replication stops at the critical time instant  $\delta^*$  when host cell activities are permanently shutdown due to the infection. This critical time was shown to decrease with increasing MOIs [15; 16; 18] or equivalently with increasing intracellular viral load [19]:

$$\delta^* = \frac{\delta_{10}}{\text{LOG}_{10} \left( \sum_k \mathbf{dna}_k \right)} \quad \text{Eq. 3.11}$$



with  $\delta_{10}$  (hpi) the critical time instant correspondent to 10 intracellular vDNA copies. The third term in equation 3.9 is the virus budding rate,  $r_{bud,k}$  (dna.cell<sup>-1</sup>.h<sup>-1</sup>):

$$r_{bud,k} = \begin{cases} k_{bud} dna_k, & \delta_{bud,low} \leq t - TOI \leq \delta_{bud,high} \\ 0, & otherwise \end{cases} \quad \text{Eq. 3.12}$$

with  $\delta_{bud,low}$  (hpi) and  $\delta_{bud,high}$  (hpi) the time of onset and halt of virus budding, respectively, and  $k_{bud}$  (h<sup>-1</sup>) the first-order virus budding constant. The transcription of foreign gene **k** into the corresponding mRNA under the control of polh promoter is defined by equation 3.13:

$$\frac{d rna_k}{dt} = r_{transc,k} - k_{DRNA,k} rna_k, \quad rna_k(0) = 0 \quad \text{Eq. 3.13}$$

with  $rna_k$  the intracellular concentration of mRNA coding for protein **k** (rna.cell<sup>-1</sup>),  $r_{transc,k}$  (rna.cell<sup>-1</sup>.h<sup>-1</sup>) the transcription rate of gene **k** and  $k_{DRNA,k}$  (h<sup>-1</sup>) the first order degradation rate. The transcription rate is as follows:

$$r_{transc,k} = \begin{cases} k_{rna,elong} \frac{dna_k}{S_k} \left( 1 - \frac{\delta^* - (t - TOI)}{\delta^* - \delta_{polh,low}} \right), & \delta_{polh,low} \leq t - TOI \leq \delta^* \\ 0, & otherwise \end{cases} \quad \text{Eq. 3.14}$$

with  $k_{rna,elong}$  (bp.h<sup>-1</sup>) the RNA polymerase elongation rate,  $S_k$  (bp) the size of gene **k** and  $\delta_{polh,low}$  (hpi) the characteristic time post infection for the onset of polh promoter transcription. Also a linear decay in vDNA transcription is considered due to the metabolic burden effect. The translation of mRNA into viral proteins is defined by equation 3.15:

$$\frac{dvp_k}{dt} = r_{transl,k}, \quad vp_k(0) = 0 \quad \text{Eq. 3.15}$$

with  $vp_k$  (μg.cell<sup>-1</sup>) the concentration of intracellular viral protein **k** and the  $r_{transl,k}$  (μg.cell<sup>-1</sup>.h<sup>-1</sup>) the corresponding synthesis rate:

$$r_{transl,k} = \begin{cases} k_{rib,elong} \frac{rna_k}{S_k} mp_k \left( 1 - \frac{\delta^* - (t - TOI)}{\delta^* - \delta_{polh,low}} \right) i_{transl}, & \delta_{polh,low} \leq t - TOI \leq \delta^* \\ 0, & otherwise \end{cases} \quad \text{Eq. 3.16}$$

with  $k_{rib,elong}$  (bp.rna<sup>-1</sup>.h<sup>-1</sup>) the ribosomal elongation rate,  $mp_k$  (μg) the molecular weight of  $vp_k$  and  $i_{transl}$  an adimensional parameter to account for the hypothesis of a maximum translation capacity. The term  $i_{transl}$  is defined here as:

$$i_{transl} = \begin{cases} 1, & L_{transl} \leq L_{transl,max} \\ \frac{L_{transl,max}}{L_{transl}}, & L_{transl} > L_{transl,max} \end{cases} \quad \text{Eq. 3.17}$$

where  $L_{transl,max}$  (rna.bp.cell<sup>-1</sup>) is the maximum translation capacity and  $L_{transl}$  (rna.bp.cell<sup>-1</sup>) the translation machinery utilization in case of all mRNA molecules being translated at the same time, defined by:

$$L_{transl} = \sum_k S_k rna_k \quad \text{Eq. 3.18}$$

### 3. CASE STUDY: PRODUCTION OF VP2

#### 3.1. Model calibration

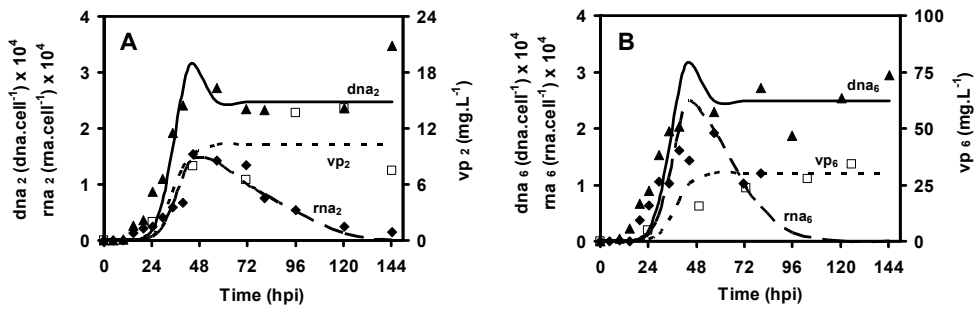
The expression of  $vp_2$  by infecting Sf-9 suspended cultures with monocistronic recombinant baculoviruses is addressed in this study. Intra and extracellular data of vDNA, mRNA and viral protein, collected over time in experiments with MOI of 5 virus.cell<sup>-1</sup> and a starting cell density  $n(0) = 10^6$  cell.ml<sup>-1</sup>, was used to calibrate the model parameters. Nonlinear fitting by the Levenberg-Marquardt method was adopted using MATLABs' functions *lsqnonlin* and *nlpaci*. The latter was used for the determination of parameters' 95% confidence limits. Table 8 compiles the calibrated parameters along with the parameters taken from the literature.

**Table 8.** Mathematical model parameters

Parameter	Value	Reference
$\mu_{max}$	$0.03 \text{ h}^{-1}$	[21]
$k_{d,int}$	$0.0008 \text{ h}^{-1}$	[16; 19; 22]
$\Delta k_d$	$0.0017 \pm 0.0003 \text{ h}^{-1}$	[19]
$\delta_{reinf}$	10 hpi	[16]
$k_a$	$7.8 \times 10^{-8} \text{ ml.cell}^{-1}.\text{h}^{-1}$	[20]
$\eta_t$	0.5	[11]
$k_{RNA}$	$0.49 \pm 0.03 \text{ h}^{-1}$	[23]
$\delta_{DNA,low}$	6 hpi	[10; 24; 25]
$\delta_{10}$	$233 \pm 7 \text{ hpi}$	[19]
$\delta_{bud,low}$	16 hpi	[15; 16]
$\delta_{bud,high}$	52 hpi	[15; 16]
$k_{bud}$	$0.07 \pm 0.02 \text{ h}^{-1}$	[23]
$k_{DRNA,2}$	$0.019 \pm 0.004 \text{ h}^{-1}$	[19]
$k_{rna,elong}$	$325 \pm 69 \text{ bp.h}^{-1}$	[23]
$S_2$	2690 bp	[19]
$\delta_{polh,low}$	15 hpi	[14; 19]
$k_{rib,elong}$	$5.5 \pm 0.6 \times 10^6 \text{ bp.rna}^{-1}.\text{h}^{-1}$	[23]
$mp_2$	$1.69 \times 10^{-13} \mu\text{g}$	[26]
$L_{transl,max}$	$1.1 \pm 0.3 \times 10^7 \text{ rna.bp.cell}^{-1}$	[23]

Figure 15A compares model and experimental data of intracellular **dna<sub>2</sub>**, **rna<sub>2</sub>** and of total (intracellular plus extracellular) **vp<sub>2</sub>**. This data shows an initial “lag” phase corresponding to early-gene expression. Although baculoviruses encode their own RNA polymerase and transcriptional factors, activators and enhancers, it is during this phase that they require the host RNA polymerase for the transcription of early genes essential for initiation of viral DNA replication such as the DNA polymerase and the anti-apoptotic p35 gene, amongst others.

This phase is followed by an exponential **dna<sub>2</sub>** increase with a characteristic vDNA replication rate of  $0.49 \text{ h}^{-1}$ . Budding of replicated viruses is known to commence around 16 hpi and enduring for almost 36 h. Virus budding is the direct cause of the decrease in intracellular **dna<sub>2</sub>** shown in figure 15A. The **dna<sub>2</sub>** continues to replicate until a characteristic time instant, around 48 hpi, after which the host cell machinery is no longer capable to express viral proteins - the metabolic burden effect.



**Figure 15.** Process kinetics of vDNA replication, mRNA synthesis and viral protein production for the genes coding for  $\text{vp}_2$  (A) and  $\text{vp}_6$  (B). The lines indicate model simulations: the full line (—) and dashed line (---) represents intracellular  $\text{dna}_k$  ( $\text{dna.cell}^{-1}$ ) and  $\text{rna}_k$  ( $\text{rna.cell}^{-1}$ ) respectively, whereas the dotted line (....) indicates total  $\text{vp}_k$  concentration ( $\text{mg.L}^{-1}$ ). Experimental data is denoted by closed triangles ( $\blacktriangle$ ) and diamonds ( $\blacklozenge$ ) in the cases of intracellular  $\text{dna}_k$  and  $\text{rna}_k$  respectively, and open squares ( $\square$ ) for total  $\text{vp}_k$  concentration.

Figure 15B shows model simulations for a viral protein with a different gene size, rotavirus  $\text{vp}_6$ . The production of rotavirus protein  $\text{vp}_6$  is achieved by infecting Sf-9 cells with monocistronic recombinant baculovirus a MOI of 5 virus.cell $^{-1}$  and a starting cell density  $n(0) = 10^6$  cell.ml $^{-1}$ . The gene coding for this protein has  $S_6 = 1356$  bp. All model parameters were kept in this simulation except for the new value of the gene size and the mRNA degradation rate [19; 27]. It can be seen that viral protein prediction is within the range of the measurements, given the high error associated with some of the techniques used. Also, the vDNA and mRNA simulations are in agreement with the experimental data.

### 3.2. Stochastic simulation

The model equations 3.1 to 3.18 is a large stochastic ordinary dynamical system. With a population number  $\text{npop} = 1000$  (the number adopted in all calculations in this study), 5000 nonlinear ordinary differential equations (ODE) must be solved over a time window of 0-100 h. Several simulation trials should also be repeated in order to extract the average behaviour of 100

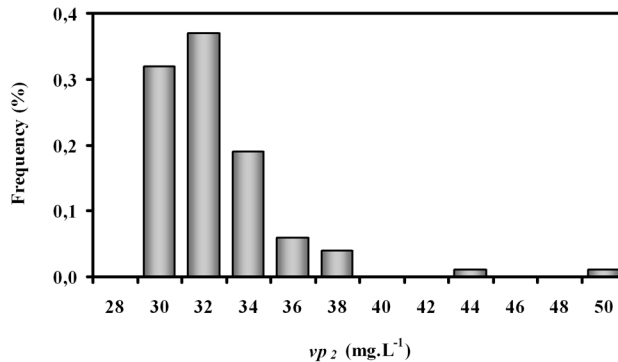
the system. To solve this large system at bearable computation time, an explicit discrete time formulation was adopted. The value of  $\mathbf{X}_k$  in equation 3.5 was obtained by a random number generator according to the Poisson distribution equations 3.6 and 3.7 (MATLABs' *poissrnd* function) over a fixed time step,  $\mathbf{T}(\mathbf{h})$ , whereas the dynamics over this time step were computed either by the direct Euler method or by a 4<sup>th</sup>/5<sup>th</sup> order Runge-Kutta solver (MATLABs' *ode45* function). Table 9 shows the computation time along with the final prediction of  $\mathbf{vp}_2$  and infected cell concentrations for different time steps. For time steps between 0.25 h and 0.01 h, the Euler method produces fairly constant results. The increase of the population number from 100 to 10000 does not improve significantly the accuracy. The *ode45* solver results in a much higher computation times. The *ode45* results converge with those of the Euler method as the time step is decreased. From the analysis above, the Euler method with  $\mathbf{T} = 0.1$  h and  $\mathbf{npop} = 1000$  was taken as a reasonable compromise between accuracy and computation time. These settings were adopted in all calculations presented in Table 9.

**Table 9.** Computation time and accuracy.

Method	dt (h)	npop	$\mathbf{vp}_2$ (mg.L <sup>-1</sup> )	$\mathbf{n}_i$ (cell.ml <sup>-1</sup> )	Cpu (s)
Euler	2	100	52.4	2997285	2
	1	100	68.3	2945690	3.75
	0.5	100	74.7	2865203	7.4
	0.25	100	80.5	2880747	14.3
	0.1	100	82.1	2845599	35.4
	0.01	100	80.7	2800949	372.2
	0.1	1000	79.9	2772313	286.5
	0.1	10000	79.0	2741385	2808.8
ode45	2	100	78.9	2568988	51.9
	1	100	79.4	2609264	98.9
	0.5	100	80.2	2678395	191.4
	0.25	100	78.6	2674363	374.5
	0.1	100	79.9	2735712	930.2
	0.05	100	84.1	2847619	1837.6

To illustrate the effect of the stochastic infection process equations 3.5 to 3.7, the statistics of final  $\mathbf{vp}_2$  productions for 100 trials and a MOI = 0.1

virus.cell<sup>-1</sup> are presented in the histogram of Figure 16.

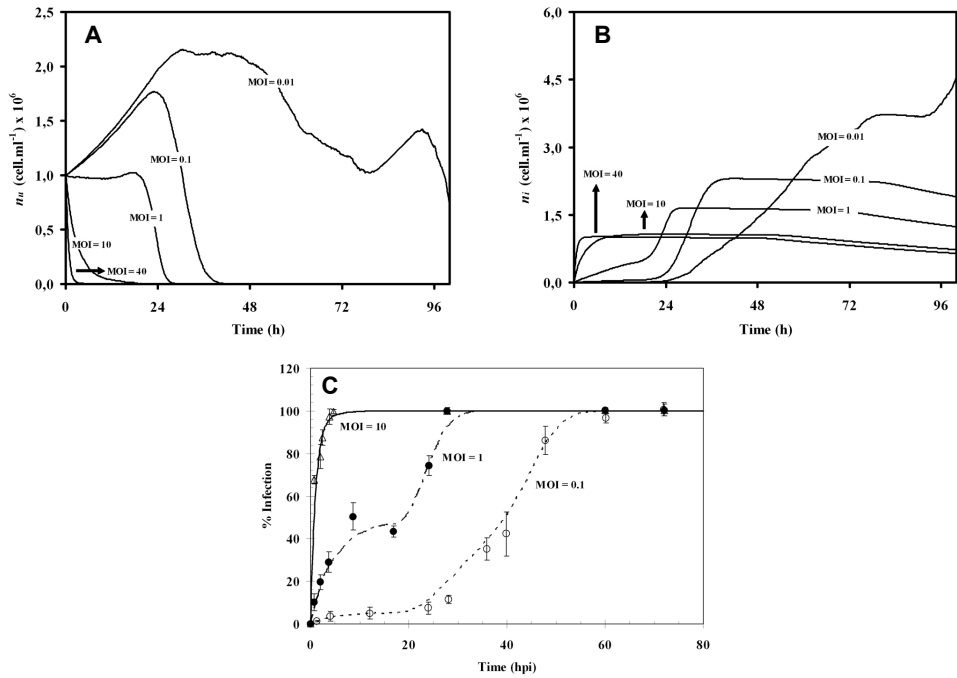


**Figure 16.** vp<sub>2</sub> distribution in the stochastic infection process. Frequency histogram of total vp<sub>2</sub> concentration (mg.L<sup>-1</sup>) for N<sub>trials</sub> = 100 and MOI = 0.1 virus.cell<sup>-1</sup>.

### 3.3. Population dynamics

The dynamics of healthy,  $n_u$  (cell.ml<sup>-1</sup>), and infected,  $n_i$  (cell.ml<sup>-1</sup>), cells for MOI = 0.01, 0.1, 1, 10 and 40 virus.cell<sup>-1</sup> are shown in Figures 17A and 17B respectively. For high MOIs, 100% infection occurs within the first 3 hours. The total cell density practically does not increase above the initial cell density. On the other extreme, for MOIs in the range 0.1-0.01 virus.cell<sup>-1</sup> a high percentage of cells remain healthy and growing in the first 36 h. It is with virus budding at 16-36 hpi, starting the second round of infection, that a major increase of the infected cell population occurs (see the case of MOI = 0.1 virus.cell<sup>-1</sup>). In the second round of infection, all healthy cells are finally infected at MOI = 0.1 virus.cell<sup>-1</sup>. In the case of MOI = 0.01 virus.cell<sup>-1</sup> a residual percentage of healthy cells remains at end of the run.

Figure 17C compares model simulations with experimental data of Hu and Bentley (2001) for MOIs of 0.1, 1 and 10 virus.cell<sup>-1</sup>. It can be observed an excellent agreement between model simulations and experimental percentage of infected population. This data is consistent with a saturation of the budding rate to 1.5 dna<sub>k</sub>.cell<sup>-1</sup>.h<sup>-1</sup>.



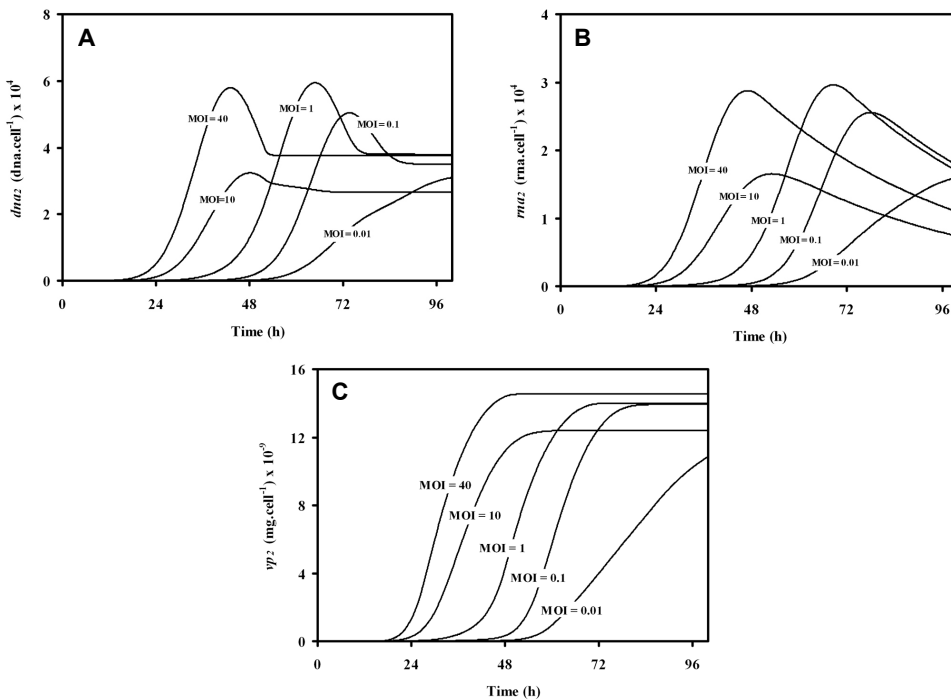
**Figure 17.** Kinetic profile of uninfected  $n_u$  (cell.mL<sup>-1</sup>) (A) and infected  $n_i$  (cell.mL<sup>-1</sup>) (B) cells population. Process simulations were performed using the stochastic infection model for MOI = 0.01, 0.1, 1, 10 and 40 virus.cell<sup>-1</sup>, and  $N_{\text{trials}} = 30$ . Infection efficiency vs time, comparison of model predictions with Hu and Bentley (2000) data (C). Model predictions are indicated by lines, the numbers next to the lines indicate the MOI used. Experimental data is denoted by symbols: MOI = 0.1 (○), MOI = 1 (●) and MOI = 10 virus.cell<sup>-1</sup> (△).

### 3.4. Intracellular dynamics

Figure 18 shows the dynamics of intracellular **dna<sub>2</sub>**, **rna<sub>2</sub>** and **vp<sub>2</sub>** in the host *Sf-9* cells after infection for different MOIs. The lines represent the average values over all infected populations.

By increasing the MOI, the maximum vDNA contents is reached sooner but it does not necessarily translates into higher intracellular vDNA contents. The maximum intracellular vDNA content is a trade-off between two different factors: high MOI represents more ‘seeds’ for vDNA replication but reduces the time window for vDNA replication. The other important factor is

that for low MOIs virus budding is extended over a long period of time, which will continuously infect and re-infect the existing host cells. For  $\text{MOI} = 10$  and  $40 \text{ virus.cell}^{-1}$ , only the first round of infection occurs while for  $\text{MOI} = 1, 0.1$  and  $0.01 \text{ virus.cell}^{-1}$  the second round of infection is also active. For low MOIs, it is only after the second round of infection that the increase in intracellular vDNA becomes visible.



**Figure 18.** Intracellular  $\text{dna}_2$  (dna.cell<sup>-1</sup>) (A),  $\text{rna}_2$  (rna.cell<sup>-1</sup>) (B) and  $\text{vp}_2$  (mg.cell<sup>-1</sup>) (C) profile for  $\text{MOI} = 0.01, 0.1, 1, 10$  and  $40 \text{ virus.cell}^{-1}$ . Process simulations were performed using the stochastic infection model for  $N_{\text{trials}} = 30$ .

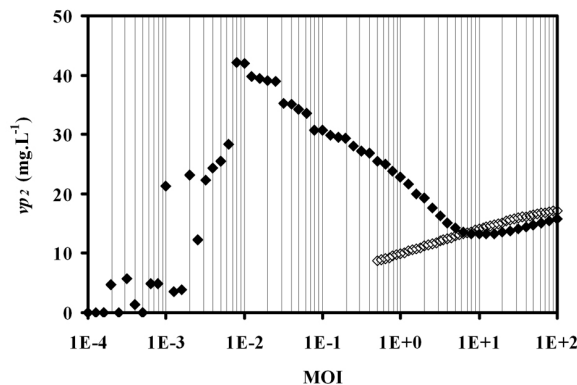
As expected, the intracellular  $\text{rna}_2$  follows closely the vDNA patterns since a direct first order transcription law was applied. The  $\text{rna}_2$  is produced in the very late phase and, as a rule, it can be stated that high  $\text{dna}_2$  loads result in high  $\text{rna}_2$  levels. The  $\text{rna}_2$  and  $\text{dna}_2$  synthesis halts at approximately the same time. The decay observed in  $\text{rna}_2$  is due to degradation. Finally, the intracellular  $\text{vp}_2$  dynamics follows closely the  $\text{dna}_2$  and  $\text{rna}_2$  dynamics.



Saturation is however observed for high MOIs. This means that it does not bring any advantage to increase the MOI to very high levels as the translation machinery saturates, thus resulting in no improvement in protein expression.

### 3.5. Productivity assessment

The final process productivity is a compromise between high intracellular  $vp_2$ , obtained at high MOIs, and high infected cell density, obtained at low MOIs. An optimal midterm MOI exists that maximises process productivities. Figure 19 plots the final  $vp_2$  production ( $\text{mg.L}^{-1}$ ) for MOIs within the range of  $10^{-4}$  to  $10^2$  virus.cell $^{-1}$  in the case of the stochastic infection model whereas in the deterministic one the MOI varies from  $10^{-0.3}$  to  $10^2$  virus.cell $^{-1}$ .



**Figure 19.** Effect of MOI on  $vp_2$  production levels. Simulations of total  $vp_2$  concentration ( $\text{mg.L}^{-1}$ ) using the stochastic and the deterministic model are represented by closed (◆) and open diamonds (◇) respectively. The MOI interval tested for the stochastic case ranges from  $10^{-4}$  to  $10^2$  virus.cell $^{-1}$  whereas in the deterministic one the MOI varies from  $10^{-0.3}$  to  $10^2$  virus.cell $^{-1}$ .  $N_{\text{trials}} = 30$ .

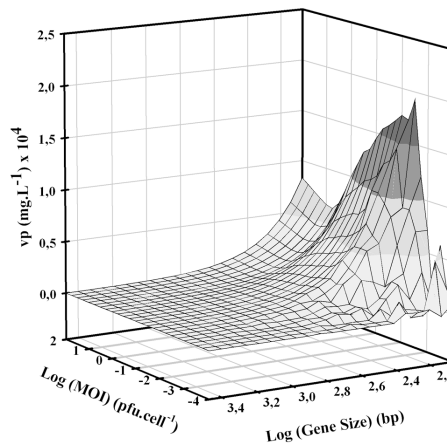
This curve shows a maximum productivity at MOI around  $10^{-2}$  virus.cell $^{-1}$  and a local minimum at MOI = 10 virus.cell $^{-1}$ . The local minimum at MOI = 10 virus.cell $^{-1}$  can be attributed to the following two peculiar factors: the

MOI is sufficiently high for 100% infection in the first infection cycle but insufficient for the saturation of the translation machinery. By the time that virus budding commences, re-infection is no longer possible, thus the intracellular protein expression remains below the saturation level. For higher MOIs than 10 virus.cell<sup>-1</sup>, no significant improvement in **vp<sub>2</sub>** production is observed due to the protein expression saturation. Moreover, the infection at such high MOIs becomes essentially synchronous and instantaneous, thus both deterministic and stochastic models are expected to converge to the same solution.

These results are in agreement with the data of Hu and Bentley (2000) who showed that **vp<sub>2</sub>** production increases as the MOI decreases from 10 to 0.1 virus.cell<sup>-1</sup>. The **vp<sub>2</sub>** production at MOI = 10 virus.cell<sup>-1</sup> was measured to be slightly higher than the productivity at MOI = 1 virus.cell<sup>-1</sup> suggesting that the local minimum is somewhere between 1 and 10 virus.cell<sup>-1</sup> but probably close to 10 virus.cell<sup>-1</sup>. The dynamic patterns and characteristic time windows are remarkably similar in both studies. The final **vp<sub>2</sub>** values are however not comparable due to different cell growth conditions. Also Power *et al.* (1994) [28] determined experimentally that the production of  $\beta$ -galactosidase in the AcMNPV/Sf-9 system decreases with increasing MOI from 0.01 to 10 virus.cell<sup>-1</sup>. Between MOI of 10 and 100 virus.cell<sup>-1</sup> no significant increase in the productivity was observed supporting the possibility of a local minimum around MOI of 10 virus.cell<sup>-1</sup>.

Figure 20 evaluates the effect of the gene size and MOI on the final protein production. This simulation shows that the increase in gene size results in the decrease of the protein productivity as showed elsewhere [6]. Interestingly, the optimal MOI seems to be independent of the gene size. This result is rational because the recombinant gene size does not affect either the infection kinetics or the vDNA replication rate or the budding of

progeny virus. It only affects the transcription and translation mechanisms, meaning that it only impacts on the intracellular accumulation of protein.



**Figure 20.** Gene size and MOI effect on final recombinant protein production. Surface plot of viral protein production ( $\text{mg.L}^{-1}$ ) for a MOI range between  $10^{-4}$  to  $10^2$  virus.cell<sup>-1</sup> and gene sizes varying within  $10^2$  to  $10^{3.5}$  bp, and  $N_{\text{trials}} = 30$ .

#### 4. CONCLUSIONS AND FUTURE PERSPECTIVES

In this paper, a stochastic simulation model of the baculovirus/insect cells system is proposed. The model combines an explicit stochastic infection process with mass action kinetics describing the dynamics of intracellular vDNA, mRNA and foreign protein. By combining these two factors, the proposed model represents an advance over previous modelling studies for the baculovirus/insect cells system. The model was shown to describe the typical behaviour of experimental data. The study clearly shows that a stochastic modelling framework is essential for the optimisation of protein expression in this system. The results confirm that the optimal MOI is typically low, in agreement with many experimental observations and other modelling studies [16]. This model could be particularly useful in designing co-infection experiments for the production of virus-like particles (VLPs)

since the statistics of infection becomes a crucial factor as one must guarantee that the infected cells carry all genes required for VLP assembly.

## 5. ACKNOWLEDGEMENTS

The authors wish to thank Helena Vieira, Catarina Estêvão and Maria Candida Mellado for thoughtful discussions. This work was supported by European Commission project (QLRT-2001-01249) and the Portuguese Fundação para a Ciência e Tecnologia (POCTI/BIO/55975/2004 and SFRH/BD/21910/2005).

## 6. REFERENCES

1. Arif, B. 1986. The Molecular Biology of Baculoviruses. In: W.D.a.P. Bohm (Ed), Springer, Berlin, pp. 31-50.
2. Cruz, P.E., Cunha, A., Peixoto, C.C., Clemente, J., Moreira, J.L., Carrondo, M.J., **1998**. Optimization of the production of virus-like particles in insect cells. *Biotechnol Bioeng.* 60 (4): 408-418.
3. Haas, R., Nielsen, L.K., **2005**. A physiological product-release model for baculovirus infected insect cells. *Biotechnol Bioeng.* 91 (6): 768-72.
4. Jang, J.D., Sanderson, C.S., Chan, L.C.L., Barford, J.P., Reid, S., **2000**. Structured modeling of recombinant protein production in batch and fed-batch culture of baculovirus-infected insect cells. *Cytotechnology.* 34 (1-2): 71-82.
5. Mortola, E., Roy, P., **2004**. Efficient assembly and release of SARS coronavirus-like particles by a heterologous expression system. *FEBS Lett.* 576 (1-2): 174-8.
6. Palomares, L.A., Lopez, S., Ramirez, O.T., **2002**. Strategies for manipulating the relative concentration of recombinant rotavirus structural proteins during simultaneous production by insect cells. *Biotechnol Bioeng.* 78 (6): 635-644.
7. Ramachandran, A., Bashyam, M.D., Viswanathan, P., Ghosh, S., Kumar, M.S., E., H., **2001**. The bountiful and baffling baculovirus: The story of polyhedrin transcription. *Current Science.* 81 (8): 998-1010.
8. Vieira, H.L., Estevao, C., Roldao, A., Peixoto, C.C., Sousa, M.F., Cruz, P.E., Carrondo, M.J., Alves, P.M., **2005**. Triple layered rotavirus VLP production: kinetics of vector replication, mRNA stability and recombinant protein production. *J Biotechnol.* 120 (1): 72-82.
9. Summers, M.D., Smith, G.E., 1987. A manual of methods for baculovirus vectors and insect cell culture procedures.
10. Blissard, G.W., Rohrmann, G.F., **1990**. Baculovirus diversity and molecular biology. *Annu Rev Entomol.* 35 127-155.
11. Dee, K.U., Shuler, M.L., **1997**. A mathematical model of the trafficking of acid-dependent enveloped viruses: Application to the binding, uptake, and nuclear accumulation of baculovirus. *Biotechnol Bioeng.* 54 (5): 468-490.
12. Power, J.F., Reid, S., Greenfield, P.F., Nielsen, L.K., **1996**. The kinetics of baculovirus adsorption to insect cells in suspension culture. *Cytotechnology.* 21 (2): 155-163.

13. Volkman, L.E., Goldsmith, P.A., **1985**. Mechanism of neutralization of budded *Autographa Californica* Nuclear Polyhedrosis Virus by a monoclonal antibody: Inhibition of entry by adsorptive endocytosis. *Virology*. 143 185-195.
14. Possee, R.D., King, L.A., **1992**. *The Baculovirus Expression System: A laboratory guide*, London.
15. Hu, Y.C., Bentley, W.E., **2001**. Effect of MOI ratio on the composition and yield of chimeric infectious bursal disease virus-like particles by baculovirus co-infection: deterministic predictions and experimental results. *Biotechniques*. 75 (1): 104-119.
16. Hu, Y.-C., Bentley, W.E., **2000**. A kinetic and statistical-thermodynamic model for baculovirus infection and virus-like particle assembly in suspended insect cells. *Chem Eng Sci*. 55 3991-4008.
17. Fischer, T.K., Bresee, J.S., Glass, R.I., **2004**. Rotavirus vaccines and the prevention of hospital-acquired diarrhea in children. *Vaccine*. 22 Suppl 1 S49-54.
18. Licari, P., Bailey, J.E., **1992**. Modeling the population dynamics of baculovirus-infected insect cells: Optimizing infection strategies for enhanced recombinant protein yields. *Biotechnol Bioeng*. 39 (4): 432-441.
19. Roldão, A., Vieira, H.L., Charpilienne, A., Poncet, D., Roy, P., Carrondo, M.J., Alves, P.M., Oliveira, R., **2007**. Modeling rotavirus-like particles production in a baculovirus expression vector system: Infection kinetics, baculovirus DNA replication, mRNA synthesis and protein production. *J Biotechnol*. 128 (4): 875-894.
20. Dee, K.U., Shuler, M.L., **1997**. Optimization of an assay for baculovirus titer and design of regimens for the synchronous infection of insect cells. *Biotechnol Prog*. 13 (1): 14-24.
21. Wang, M.-Y., Vakharia, V., Bentley, W.E., **1993**. Expression of epoxide hydrolase in insect cells: A focus on the infected cell. *Biotechnol Bioeng*. 42 (2): 240 - 246.
22. Dalal, N.G., Bentley, W.E., **1999**. Mathematical characterization of insect cell (Sf-9) death in suspended culture. *Biotechnol Lett*. 21 325-329.
23. Roldão, A., Carrondo, M.J.T., Alves, P.M., Oliveira, R., **2008**. Stochastic simulation of protein expression in the baculovirus/insect cells system. *Computers & Chemical Engineering*. 32 (1-2): 68-77.
24. Carstens, E.B., Tjia, S.T., Doerfler, W., **1979**. Infection of *Spodoptera frugiperda* cells with *Autographa californica* nuclear polyhedrosis virus I. Synthesis of intracellular proteins after virus infection. *Virology*. 99 (2): 386-98.
25. Tjia, S.T., Carstens, E.B., Doerfler, W., **1979**. Infection of *Spodoptera frugiperda* cells with *Autographa californica* nuclear polyhedrosis virus II. The viral DNA and the kinetics of its replication. *Virology*. 99 (2): 399-409.
26. Labbe, M., Charpilienne, A., Crawford, S.E., Estes, M.K., Cohen, J., **1991**. Expression of rotavirus VP2 produces empty corelike particles. *J Virol*. 65 (6): 2946-52.
27. Roldão, A., Vieira, H.L.A., Alves, P.M., Oliveira, R., Carrondo, M.J.T., **2006**. Intracellular dynamics in rotavirus-like particles production: Evaluation of multigene and monocistronic infection strategies. *Process Biochemistry*. 41 (10): 2188-2199.
28. Power, J.F., Reid, S., Radford, K.M., Greenfield, P.F., Nielsen, L.K., **1994**. Modeling and optimization of the baculovirus expression vector system in batch suspension culture. *Biotechnol Bioeng*. 44 (6): 710-719.



# Chapter 4

## Thermodynamic-equilibrium based RLP assembly

This chapter was based on the following manuscript:

António Roldão, Maria Candida M. Mellado, J. C. Lima, Manuel J.T. Carrondo, Paula M. Alves and R. Oliveira (2010). "A Thermodynamic Equilibrium-Based Model for *in vitro* Analysis and Design of Virus-Like Particles (VLP) Assembly: The Case of Rotavirus VLP", *Biophysical Journal* (under review).

## Summary

In this study, a thermodynamic equilibrium-based model describes for the first time the *in vitro* assembly of rotavirus-like particles. Assembly and disassembly data collected at different *in vitro* physicochemical conditions is analyzed at the aid of this model. Equilibrium calculations show that under favourable conditions particles distribution are dominated by structural subunits and completely built icosahedra, while other intermediates are present only at residual concentrations. There is a minimum concentration of structural subunits that ensures measurable particle formation, i.e. the pseudo-critical concentration level. This threshold concentration value is shown to be coupled to the Gibbs free energy of subunit association. For low Gibbs free energies, thermodynamics maximizes the formation of intact icosahedra whenever proteins are provided at the exact stoichiometric ratio. At high Gibbs free energies, there is a translocation of proteins pseudo-critical concentrations; so, optimum concentrations are no longer coincident to the particle stoichiometric composition. The dependency between assembly efficiency and Gibbs free energy of subunit association is not linear. For values of Gibbs free energy of subunit association below  $-2 \text{ kcal.mol}^{-1}$ , changes in Gibbs free energy have no significant effect on the assembly efficiency. Above this threshold, a crisp transition from intact icosahedra to unassembled subunits is observed. In the end, this work provides a computational framework for the analysis of multilayered particles assembly and clearly demonstrates that particle formation can be controlled through the oriented manipulation of the Gibbs free energy of subunit association.



## CONTENTS

<b>1. Introduction .....</b>	<b>114</b>
<b>2. Materials and methods .....</b>	<b>116</b>
2.1. RLPs production, purification and analysis .....	116
2.2. <i>In vitro</i> assembly and disassembly assays .....	117
<b>3. A model for rotavirus-like particle assembly.....</b>	<b>118</b>
3.1. Zlotnick's equilibrium framework .....	120
3.2. Application to the self-assembly of RLP .....	122
3.2.1. Formation of the vp2 layer .....	123
3.2.2. Formation of the vp6 layer .....	125
3.2.3. Formation of the vp7 layer .....	129
<b>4. Results and discussion .....</b>	<b>132</b>
4.1. Pseudo-critical protein concentration .....	132
4.2. The effect of the number of association contacts.....	133
4.3. Simulation of RLP assembly .....	134
4.4. <i>In vitro</i> RLPs assembly experiments .....	137
4.5. Temperature dependency .....	138
4.6. Interpretation of <i>in vitro</i> disassembly data .....	141
4.7. Design of optimal thermodynamic conditions for global RLP assembly ....	142
<b>5. Conclusion.....</b>	<b>145</b>
<b>6. Acknowledgments .....</b>	<b>146</b>
<b>7. References.....</b>	<b>146</b>

## 1. INTRODUCTION

*In vivo* assembly of virus-like particles (VLP) results many times in predominantly malformed structures because the host biology does not provide the adequate thermodynamic environment for the correct assembly of such particles. A well known example is the prophylactic human papillomavirus (HPV)-VLP vaccine developed by Merck & Co., Inc. (Gardasil™) where *in vitro* disassembly and reassembly steps are essential to maximise the production of the correctly assembled VLP from the L1 proteins of HPV types 6, 11, 16 and 18.

The *in vivo* assembly of complex multilayered particles, as the rotavirus-like particles (RLP) model studied here, is even more challenging. RLP are spherical shaped particles composed by three viral proteins (vp) of rotavirus arranged in a triple layered structure: vp2 - 102.7 kDa (innermost layer) [1]; vp6 - 44.9 kDa (middle shell) [2]; vp7 - 37.2 kDa (outer layer) [2]. *In vivo* assembly efficiency data shows that contaminants such as trimers of vp6 and vp7, vp6 tube-like structures, single-layered vp2 particles, double layered particles of vp2 and vp6 (DLP) or RLP lacking one or more subunits represent almost 88% of the total mass of proteins expressed [3]. Unfortunately, few experimentally supported information exists of intra and inter-protein binding energies, rates and orders for assembly reactions that could enlighten the causes of particle malformation. A recent study of *in vitro* assembly/disassembly of RLPs explored how physicochemical parameters affect the formation and stability of RLPs [4]. This study showed that protein macrostructures are highly dependent on the pH, ionic strength and temperature. However, designing the ideal physicochemical environment for VLP assembly is nontrivial as the assembly pathway is complex involving many intermediate species. Mathematical models built on stability principles of protein macrostructures and respective assembly

pathways may be of great value to better understand and design *in vitro* VLP assembly experiments.

During the last decade, many mathematical models have described viral capsid assembly based on molecular dynamics [5], thermodynamics as self-organization of discs on a sphere [6], combinatorial optimization studies [7], “local rules” approach [8], stochastic icosahedral capsid growth [9] and protein aggregation by nucleated polymerization [10]. Others have used the viral tilling theory (VTT) to describe the assembly of structural subunits formed from pentamers that differ by the structure of the inter-subunit bonds that surround them, by combining the information on local environments around pentamers with the thermodynamic-equilibrium model of Zlotnick [11]. The major drawback in VTT is that distinct configurations may be energetically preferred when attaching a further building block; this gives rise to a tree of assembly pathways rather than a single assembly pathway as in Zlotnick’s model, significantly complicating the assembly process. The simplest way to describe the self-assembly of virus capsids is using Zlotnick’s thermodynamic-equilibrium formalism where the concentration of structural subunits and the Gibbs free energy of subunit association ( $\Delta G^0_n$ ) are the key variables that control particle assembly [12-16].

In this study, the self-assembly of RLP was modelled using Zlotnick’s thermodynamic-equilibrium method. A hierarchical macroassembly model is postulated assuming that the anchorage of outer layer proteins is only possible upon the complete formation of the inner layers. The effect of the initial concentrations of protein structural subunits and  $\Delta G^0_n$  on SLP assembly efficiency was analysed. The distribution dynamics of RLP intermediate structures with  $\Delta G^0_n$  was also investigated; scenarios of protein limitation and excess are explored. The dependency of particle

formation and stability with the Gibbs free energy of vp7 subunit association was explored by comparing *in silico* model simulations with *in vitro* assembly and disassembly experiments. Thermodynamic- and stoichiometric-based assembly efficiencies were estimated at different protein concentrations and Gibbs free energy of subunit association to assess the controllability for RLPs assembly.

## 2. MATERIALS AND METHODS

### 2.1. RLPs production, purification and analysis

The production of RLPs, DLPs and vp7 have been described in detail elsewhere [17, 18]. Briefly, for RLP production, *Spodoptera frugiperda* Sf-9 cells were infected at a cell concentration of  $3 \times 10^6$  cell.ml<sup>-1</sup> with three monocistronic recombinant baculoviruses enclosing the genes coding for vp2 fused to GFP (BacRF2A-GFP) [1], vp6 (BacVP6C) [19] and vp7 (BacRF7), all of bovine rotavirus strain RF. The multiplicity of infection (MOI), i.e., number of virus per cell used was 3 (vp2) + 3 (vp6) + 3 (vp7). For DLP production, Sf-9 cells were infected at a cell concentration of  $3 \times 10^6$  cell.ml<sup>-1</sup> with two monocistronic recombinant baculoviruses coding for vp2 fused to GFP (BacRF2A-GFP) and vp6 (BacVP6C); the MOI used was 3 (vp2) + 3 (vp6). For vp7 production, Sf-9 cells were infected at a cell concentration of  $1 \times 10^6$  cell.ml<sup>-1</sup> with monocistronic recombinant baculoviruses coding for vp7 (BacRF7); the MOI used was 5 virus.cell<sup>-1</sup>. All productions were performed in a 2 L stirred tank bioreactor (Sartorius-Stedim Biotech GmbH, Melsungen, Germany). The glass bioreactor was equipped with two Rushton turbines (standard geometry), a water recirculation jacket to control the temperature and a sparger for gas supply; culture temperature and gas flow were kept constant at 27 °C and 0.01 vvm

(gas volume per culture volume per minute), respectively. The percentage of dissolved oxygen in the culture was controlled at 30 % with the agitation rate (70 to 250 rpm) and the concentration of oxygen in the gas mixture (0 to 100%) (control cascade). The pH was monitored but not controlled since it did not vary significantly throughout the culture time. The abovementioned conditions were previously optimized [20].

The purification of RLPs, DLPs and vp7 was performed using *in-house* developed methods consisting in depth filtration, ultrafiltration, affinity-, size exclusion- and anion exchange-chromatography as stepwise unit operations [18, 21]. The total protein content in purified RLPs, DLPs and vp7 was determined using the BCA protein quantification assay kit (96-well plate protocol) (Pierce, Rockford, US) [17, 18].

The analysis of intact RLPs and DLPs in *in vitro* assembly and disassembly experiments was performed by capillary zone electrophoresis [4].

## 2.2. *In vitro* assembly and disassembly assays

As described in [4], *in vitro* RLPs assembly experiments using purified DLP and vp7 were carried out in TNC buffer at pH 5.5, 25 °C, NaCl 0.1 M and  $\text{Ca}^{2+}$  1 mM. The initial concentrations of DLP and vp7 were 10 pM and 7.8 nM, respectively. The rationale for this choice is providing DLPs and vp7 at concentrations that obey to the RLP stoichiometric composition (vp2:vp6:vp7 - 120:780:780 molecules). The effect of physicochemical parameters on the Gibbs free energy of vp7 subunit association and assembly efficiency was investigated. Higher pH (8), temperature (35 °C), concentration of NaCl (0.5 M) or  $\text{Ca}^{2+}$  (5 mM) were tested (see Table 13). Once solution reaches equilibrium, assembly efficiency was calculated as:

$$\text{Ass}_{\text{eff}} = 1 - \frac{[\text{DLP}]_e}{[\text{DLP}]_0} \quad \text{Eq. 4.1}$$

with  $[\text{DLP}]_0$  the initial concentration of DLP and  $[\text{DLP}]_e$  the concentration of DLP at equilibrium.

*In vitro* RLP disassembly experiments were performed with chelating agents (EDTA or EGTA at 1 mM and 2 mM in D-PBS and TNC) at 35 °C. The composition of D-PBS and TNC buffers was as follows: D-PBS (137 mM NaCl, pH 7.4, 0.9 mM  $\text{CaCl}_2$  and 0.493 mM  $\text{MgCl}_2$ ); TNC (20 mM Tris-HCl, 100 mM NaCl, pH 7.3 and 0.9 mM  $\text{CaCl}_2$ ). The initial concentration of RLP was 7 nM. In addition, *in vitro* RLP spontaneous disassembly, at 4 °C, was followed for different storage times (4, 12 and 15 months).

The disassembly efficiency was calculated as follows:

$$\text{Dis}_{\text{eff}} = 1 - \frac{[\text{RLP}]_e}{[\text{RLP}]_0} \quad \text{Eq. 4.2}$$

with  $[\text{RLP}]_0$  the initial concentration of RLP and  $[\text{RLP}]_e$  the concentration of RLP at equilibrium.

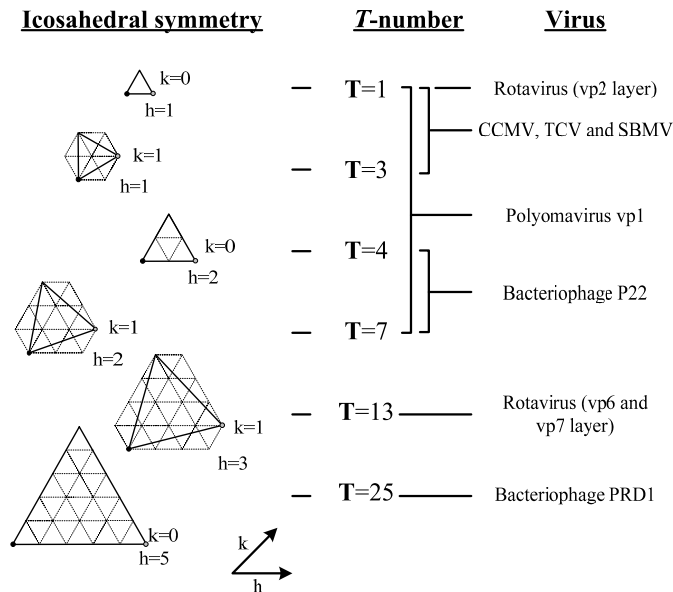
The reassembly of the particles previously disassembled with EDTA or EGTA 2 mM was studied using dialysis; the conditions used for the reassembly were 5 mM of  $\text{Ca}^{2+}$  and 25 °C.

These data allowed estimating the Gibbs free energies of vp7 subunit association to DLP by fitting the model described next.

### 3. A MODEL FOR ROTAVIRUS-LIKE PARTICLE ASSEMBLY

The self-assembly of many spherical viruses provide excellent models to study the assembly of RLPs. Bacteriophage P22 and PRD1, cowpea chlorotic mottle virus, turnip crinkle virus, southern bean mosaic virus or

polyomavirus are some examples [22-29]. These viruses possess the same icosahedral geometry as RLPs, being composed by 20 structural subunits (icosahedron faces), each an equilateral triangle with a specific  $T$ -number of identical proteins *per* asymmetric unit (Figure 21). More importantly, the build-up of these viruses' capsids is well known. These mechanisms can be extended to the stepwise assembly of multilayered particles using the Zlotnick's equilibrium method as described below.



**Figure 21.** The quasi-equivalence theory of icosahedral virus assembly. An icosahedron has 20 faces (structural subunits), each an equilateral triangle with three asymmetric units (AU). The quasi-equivalence theory of Caspar and Klug explains how a structural subunit assembles from an integral number ( $T$ ) of smaller triangles in viruses with more than 60 AU. The  $T$ -number is the number of identical proteins within each AU and is defined by  $T=h^2+hk+hk^2$  where  $h$  and  $k$  are pairs of integers. The AUs are equivalent whereas the  $T$  proteins comprising the AU are not. Structures with  $T=1$ ,  $T=3$ ,  $T=4$ ,  $T=7$ ,  $T=13$  and  $T=25$  icosahedral symmetry are shown for bacteriophage P22 and PRD1, cowpea chlorotic mottle virus (CCMV), turnip crinkle virus (TCV), southern bean mosaic virus (SBMV), polyomavirus and rotavirus.

### 3.1. Zlotnick's equilibrium framework

The self-assembly of a generic polyhedral virus can be described mathematically using the equilibrium theory proposed by Zlotnick in 1994. This modelling approach is based on the following principles:

- Structural subunits are very stable protein aggregates forming the building blocks in the assembly process;
- Structural subunits are sequentially added to the most stable intermediate until a complete layer is formed (see Table 10, 11B and 12);
- Only the most stable conformation, i.e. the arrangement with the greatest number of inter-subunit contacts, is considered;
- Assembly reactions are rapid, reaching equilibrium in milliseconds;
- Only ideal icosahedral geometry is allowed for each intermediate;
- Nucleation and cooperativity are neglected;
- The contact between adjacent subunits results in a characteristic Gibbs free energy variation,  $\Delta G_n^0$ ;
- All interactions between identical subunits have the same characteristic  $\Delta G_n^0$ , irrespective of the particle size.

The assembly of structural subunits into a complete layer involves a series of equilibrium reactions, statistical components and thermodynamic factors. The association constant,  $K_n$ , of a species  $n$ , a specific arrangement of  $m$  structural subunits, can be uncoupled into statistical components  $S_{1,n}$  and  $S_{2,n}$ , and a non-statistical association constant  $K'_n$  :



$$K_n = S_{1,n} S_{2,n} K'_n = \frac{[n]}{[n-1][1]} \quad \text{Eq.4.3}$$

The structural subunit  $m = n = 1$  is defined as the stable building block and  $n-1$  is defined as the intermediate composed of  $m-1$  subunits.  $S_{1,n}$  is defined as the ratio between two factors: the number of ways of forming a species  $n$  (from the preceding most stable intermediate,  $n-1$ , and free subunit  $n = 1$  – **build-up**) and the number of ways of dissociating the intermediate species  $n$  into the respective reactants,  $n-1$  (**build-down**).  $S_{2,n}$  describes the degeneracy of the incoming subunit in species  $n$ ; this statistical factor depends on the number of equivalent orientations that each incoming icosahedral asymmetric unit, called oligomer (cluster of proteins – the fundamental of the assembly reaction), adopts in a structural subunit  $m$ . The non-statistical association constant  $K'_n$  is a function of the number of contacts between subunits for each newly formed species  $n$ ,  $N_{c,n}$ , and  $\Delta G_n^0$ , and does not include the degeneracy of the interaction:

$$K'_n = \exp\left(-\frac{N_{c,n} \Delta G_n^0}{RT}\right) \quad \text{Eq.4.4}$$

with  $R=1.987 \text{ cal.mol}^{-1}\text{K}^{-1}$  the ideal gas constant and  $T$  the temperature (K). The concentration of each intermediate species  $n$  can be estimated by rearranging equations 4.3 and 4.4, so that:

$$[n] = [n-1][1] S_{1,n} S_{2,n} \exp\left(-\frac{N_{c,n} \Delta G_n^0}{RT}\right) = \prod_{\psi=2}^n S_{1,\psi} S_{2,\psi} \exp\left(-\frac{N_{c,\psi} \Delta G_\psi^0}{RT}\right) [1]^\psi \quad \text{Eq.4.5}$$

The concentration of each species is a function of the stable building block, species  $1$ , and  $\Delta G_n^0$ . The total protein in solution at time  $t$ ,  $\mathbf{Prot}(t)$ , can be calculated from the species concentration  $1$  to  $n$ :

$$Prot(t) = \sum_{\Psi=1}^n [\Psi] m_{\Psi} S_{2,n} z \quad \text{Eq.4.6}$$

where  $m_{\Psi}$  represents the number of structural subunits in species  $\Psi = n$  and  $z$  the number of proteins in each oligomer.

### 3.2. Application to the self-assembly of RLP

The modelling method proposed above is directly applicable to the equilibrium of  $n$  stable structural subunits forming icosahedra. Here we extend this method to the formation of the three RLP icosahedral layers. The formation of RLPs is hypothesized to be a sequence of 3 main steps: 1<sup>st</sup> step – assembly of SLPs; 2<sup>nd</sup> step – assembly of DLPs by attachment of vp6 to SLPs; 3<sup>rd</sup> step – RLPs assembly by vp7 attachment to DLPs. This sequence of steps is supported by experimental evidence that the three viral proteins are synthesized in different ribosomes: vp2 and vp7 are synthesized on free ribosomes while vp7 is synthesized on ER-associated ribosomes [30, 31]. This suggests that DLPs may assemble in the cytoplasm, bud into the endoplasmic reticulum and then acquire the outer vp7 layer [32]. The common feature to all phases is the pathway by which the end product is generated (Table 10, 11b and 12). The distinguishing factor is the attachment of the first vp6 and vp7 subunits to SLPs and DLPs, respectively.

Such an assembly model implies that particles composed by vp6 and vp7 or by vp2 and vp7 are never formed. Indeed, experiments showed that particles can assemble *in vivo* or *in vitro* without vp2, although their amounts are negligible. The same applies for most structures of vp6 or vp7, namely trimer aggregates in the form of tubules or sheets [33]. In addition, these unusual complexes are rather heterogeneous and have extremely low stability. Finally, we also assume that the assembly of viral proteins into

oligomers (dimers, trimers or pentamers) and from these into stable subunits is not affected by the presence of the reporter protein GFP.

### 3.2.1. Formation of the vp2 layer

In the absence of any other protein, recombinant vp2 form SLPs with diameters of about 45 nm and hexagonal outlines; no monomeric structures are observed [1, 34, 35]. This suggests that inner RLP layer and native rotavirus vp2 share the same unusual  $T=1$  icosahedral geometry. The assembly of such structure consists in the aggregation of vp2 molecules into dimers, subsequent formation of a stable structural subunit from 3 dimers and vp2 layer construction by the sequential addition of structural subunits to the most stable assembly intermediates. The formation of structural subunits is assumed to be an equilibrium process yet favoring the assembled subunits. The interaction between the three dimers in a subunit occurs at the 3-fold axes whereas the contact between consecutive subunits occurs at the 5-fold axes.

Table 10 lists the assembly intermediates and factors describing the assembly of SLPs. The Gibbs free energy of vp2 subunit association is defined by  $\Delta G^0_{\text{SLP}}$ . The subunits of SLPs are linked by strong hydrophobic interactions; thus,  $\Delta G^0_{\text{SLP}} \ll 0 \text{ kcal.mol}^{-1}$ .

Modeling parameters for this layer are:  $n=20$  - vp2 layer has icosahedral geometry with 20 structural subunits;  $S_{2,n}=3$  - each structural subunit consists of 3 icosahedral asymmetric units,  $z=2$  - each asymmetric unit is formed by 2 proteins;  $\Delta G^0_{\text{SLP}} = -4.08 \text{ kcal.mol}^{-1}$  [36, 37].

The effect of calcium concentration, pH and temperature on the Gibbs free energy of vp2 association was not considered in the model and is open for future studies.

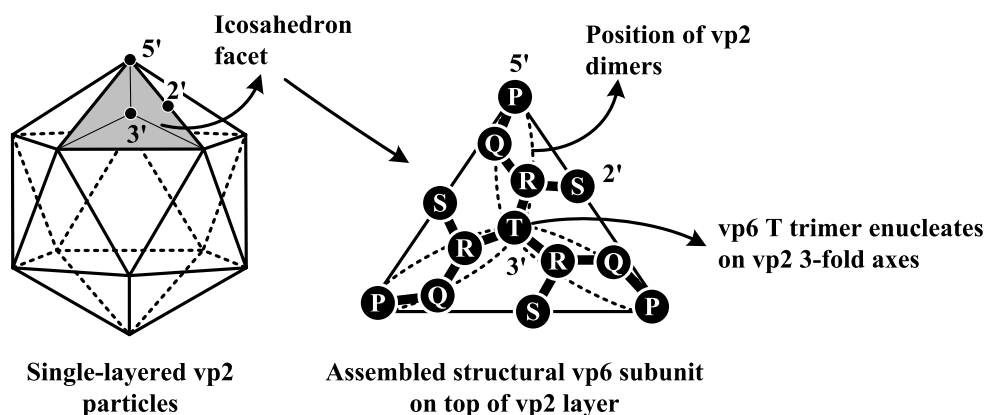
**Table 10.** Assembly intermediates and factors describing the formation of single-layered vp2 particles.

n	Model	Build		S <sub>1,n</sub>	N <sub>c,n</sub>	[n]	n	Model	Build		S <sub>1,n</sub>	N <sub>c,n</sub>	[n]
		up	down						up	down			
1		-	-	-	-	[1]	11		6	1	6/1	1	$3^{10} \times 6 \times e^{\left(-\frac{13 \times \Delta G_{SLP}^0}{RT}\right)} [1]^{11}$
2		3/2 <sup>A</sup>	1	3/2	1	$3 \times \frac{3}{2} \times e^{\left(-\frac{1 \times \Delta G_{SLP}^0}{RT}\right)} [1]^2$	12		1	4	1/4	2	$3^{11} \times \frac{3}{2} \times e^{\left(-\frac{15 \times \Delta G_{SLP}^0}{RT}\right)} [1]^2$
3		4	2	4/2	1	$3^2 \times 3 \times e^{\left(-\frac{2 \times \Delta G_{SLP}^0}{RT}\right)} [1]^3$	13		2	1	2/1	1	$3^{12} \times 3 \times e^{\left(-\frac{16 \times \Delta G_{SLP}^0}{RT}\right)} [1]^3$
4		2	2	2/2	1	$3^3 \times 3 \times e^{\left(-\frac{3 \times \Delta G_{SLP}^0}{RT}\right)} [1]^4$	14		2	2	2/2	2	$3^{13} \times 3 \times e^{\left(-\frac{18 \times \Delta G_{SLP}^0}{RT}\right)} [1]^4$
5		1	5	1/5	2	$3^4 \times \frac{3}{5} \times e^{\left(-\frac{5 \times \Delta G_{SLP}^0}{RT}\right)} [1]^5$	15		1	5	1/5	2	$3^{14} \times \frac{3}{5} \times e^{\left(-\frac{20 \times \Delta G_{SLP}^0}{RT}\right)} [1]^5$
6		5	1	5/1	1	$3^5 \times 3 \times e^{\left(-\frac{6 \times \Delta G_{SLP}^0}{RT}\right)} [1]^6$	16		5	1	5/1	1	$3^{15} \times 3 \times e^{\left(-\frac{21 \times \Delta G_{SLP}^0}{RT}\right)} [1]^6$
7		2	2	2/2	1	$3^6 \times 3 \times e^{\left(-\frac{7 \times \Delta G_{SLP}^0}{RT}\right)} [1]^7$	17		2	2	2/2	2	$3^{16} \times 3 \times e^{\left(-\frac{23 \times \Delta G_{SLP}^0}{RT}\right)} [1]^7$
8		1	2	1/2	2	$3^7 \times \frac{3}{2} \times e^{\left(-\frac{9 \times \Delta G_{SLP}^0}{RT}\right)} [1]^8$	18		2	4	2/4	2	$3^{17} \times \frac{3}{2} \times e^{\left(-\frac{25 \times \Delta G_{SLP}^0}{RT}\right)} [1]^8$
9		4	1	4/1	1	$3^8 \times 6 \times e^{\left(-\frac{10 \times \Delta G_{SLP}^0}{RT}\right)} [1]^9$	19		2	3	2/3	2	$3^{18} \times 1 \times e^{\left(-\frac{27 \times \Delta G_{SLP}^0}{RT}\right)} [1]^9$
10		1	6	1/6	2	$3^9 \times 1 \times e^{\left(-\frac{12 \times \Delta G_{SLP}^0}{RT}\right)} [1]^{10}$	20		1	20	1/20	3	$3^{19} \times \frac{1}{20} \times e^{\left(-\frac{30 \times \Delta G_{SLP}^0}{RT}\right)} [1]^{20}$

<sup>A</sup> – the build-up factor 3/2, instead of the logically factor 3, is due to the degeneracy of the asymmetric unit; **n** – represents a species composed by a specific arrangement of *m* structural subunits; **Build-up** – the number of ways of forming a species *n* from the preceding most stable intermediate, *n*-1, and free subunit *n*=1; **Build-down** – the number of ways of dissociating the intermediate species *n* into equivalent reactants, *n*-1; **S<sub>1,n</sub>** – ratio Build-up / Build-down; **N<sub>c,n</sub>** – number of contacts between subunits for each newly formed species *n*; **ΔG<sub>SLP</sub><sup>0</sup>** – Gibbs free energy of vp2 subunit association (cal.mol<sup>-1</sup>); **R** – gas constant (1.987 cal.mol<sup>-1</sup>.K<sup>-1</sup>); **T** – temperature (K); **[n]** – concentration of each intermediate species *n* (M).

### 3.2.2. Formation of the vp6 layer

Recombinant vp6 rapidly forms oligomers *in vivo*, mostly as trimeric structures. The dimeric form of the protein stabilized by disulfide bonds may exist but represent either a step in the formation of trimers or an abnormal oligomerization; monomers of vp6 are rare [35]. Upon treatment of DLPs with 1.5 M of  $\text{CaCl}_2$ , the outer layer is removed but trimers do not disaggregate. These experimental evidences support the idea of vp6 trimers being the building blocks of a RLP vp6 layer with  $T=13$  quasi-equivalent icosahedral geometry similar to the native rotavirus vp6. In such an icosahedral lattice, there are five quasi-equivalent trimers: **P**, **Q**, **R**, **S**, and **T** [38, 39]; the precise juxtaposition of these trimers varies slightly in accordance with the requirements of quasi-equivalence (Figure 22).



**Figure 22.** Diagram of the juxtaposition of vp6 trimers in relation to vp2 dimers. The precise juxtaposition of vp6 trimers varies slightly in accordance with the requirements of quasi-equivalence for the completion of a  $T=13$  lattice with 260 vp6 trimers. There are five quasi-equivalent trimers, **P**, **Q**, **R**, **S**, and **T**, with **T** trimer providing the center of the cluster located at the 3-fold axes.

The vp6 layer is formed by 20 structural subunits, each one consisting of 13 vp6 trimers. The assembly of one subunit involves the initial contact of vp6 trimer **T** with the vp2 layer at its icosahedral 3-fold axes and subsequent

interactions between trimer **T** and trimers **Q**, **R**, **S** and **P** (Table 11a and Figure 22). The assembly of vp6 layer results from the addition of structural subunits to the most stable assembly intermediates (Table 11b).

The contact that aligns the vp6 trimer **T** with the icosahedral vp2 layer is the strongest ( $\Delta G^0_{\text{vp6,T}}$ ) whereas vp6 trimer-trimer interactions within each subunit ( $\Delta G^0_{\text{vp6,PQRS}}$ ) and vp6 subunit-subunit interactions ( $\Delta G^0_{\text{DLP}}$ ) are weaker [40]. The contact between vp6 trimers **P** of adjacent subunits is ensured by strong lateral non-equivalent interactions and occurs across the quasi 2-fold axes but closer to the 3-fold axes and at the icosahedral 5-fold axes. The X-ray structure of rotavirus vp6 shows that trimer interface at the base contains mostly hydrophilic/negatively charged residues [40]. Since these residues participate in inter-layer vp2-vp6 interactions and solubilisation of vp6 layer is only accomplished using chaotropic agents, it is possible to assume for recombinant DLPs that

$$\Delta G^0_{\text{DLP}}, \Delta G^0_{\text{vp6,T}} \text{ and } \Delta G^0_{\text{vp6,PQRS}} \ll 0 \text{ kcal.mol}^{-1}$$

In virtue of model simplification and computation time, the assembly of vp6 structural subunits was lumped; hence, the Gibbs free energy of vp6 subunit association can be defined as

$$\Delta G^0_{\text{DLP}} = \Delta G^0_{\text{vp6,T}} + 24 \times \Delta G^0_{\text{vp6,PQRS}}$$






















and  $[\text{free vp6}] = [\text{vp6}_{\text{trimers}}]^{13}$ , where  $[\text{free vp6}]$  is the concentration of unassociated vp6 subunits at equilibrium (M) and  $[\text{vp6}_{\text{trimers}}]$  the concentration of vp6 trimers, the building blocks of a vp6 structural subunits.

**Table 11a.** The assembly of the first vp6 structural subunit on top of the RLP vp2 layer.

n	Model	Build		$S_{1,n}$	$N_{c,n}$	[n]
		up	down			
20.1		20	1	20/1	1	$[20] \times 20 \times e^{\left( \frac{1 \times \Delta G_{vp6,T}^0}{RT} \right)} [vp6_{trimers}]$
20.2		3	1	3/1	2	$[20] \times 60 \times e^{\left( \frac{1 \times \Delta G_{vp6,T}^0 + 2 \times \Delta G_{vp6,PQRS}^0}{RT} \right)} [vp6_{trimers}]^2$
20.3		2	2	2/2	2	$[20] \times 60 \times e^{\left( \frac{1 \times \Delta G_{vp6,T}^0 + 4 \times \Delta G_{vp6,PQRS}^0}{RT} \right)} [vp6_{trimers}]^3$
20.4		1	3	1/3	2	$[20] \times 20 \times e^{\left( \frac{1 \times \Delta G_{vp6,T}^0 + 6 \times \Delta G_{vp6,PQRS}^0}{RT} \right)} [vp6_{trimers}]^4$
20.5		6	1	6/1	2	$[20] \times 120 \times e^{\left( \frac{1 \times \Delta G_{vp6,T}^0 + 8 \times \Delta G_{vp6,PQRS}^0}{RT} \right)} [vp6_{trimers}]^5$
20.6		5	2	5/2	2	$[20] \times 300 \times e^{\left( \frac{1 \times \Delta G_{vp6,T}^0 + 10 \times \Delta G_{vp6,PQRS}^0}{RT} \right)} [vp6_{trimers}]^6$
20.7		4	3	4/3	2	$[20] \times 400 \times e^{\left( \frac{1 \times \Delta G_{vp6,T}^0 + 12 \times \Delta G_{vp6,PQRS}^0}{RT} \right)} [vp6_{trimers}]^7$
20.8		3	4	3/4	2	$[20] \times 300 \times e^{\left( \frac{1 \times \Delta G_{vp6,T}^0 + 14 \times \Delta G_{vp6,PQRS}^0}{RT} \right)} [vp6_{trimers}]^8$
20.9		2	5	2/5	2	$[20] \times 120 \times e^{\left( \frac{1 \times \Delta G_{vp6,T}^0 + 16 \times \Delta G_{vp6,PQRS}^0}{RT} \right)} [vp6_{trimers}]^9$
20.10		1	6	1/6	2	$[20] \times 20 \times e^{\left( \frac{1 \times \Delta G_{vp6,T}^0 + 18 \times \Delta G_{vp6,PQRS}^0}{RT} \right)} [vp6_{trimers}]^{10}$
20.11		3	1	3/1	2	$[20] \times 60 \times e^{\left( \frac{1 \times \Delta G_{vp6,T}^0 + 20 \times \Delta G_{vp6,PQRS}^0}{RT} \right)} [vp6_{trimers}]^{11}$
20.12		2	2	2/2	2	$[20] \times 60 \times e^{\left( \frac{1 \times \Delta G_{vp6,T}^0 + 22 \times \Delta G_{vp6,PQRS}^0}{RT} \right)} [vp6_{trimers}]^{12}$
20.13		1	3	1/3	2	$[20] \times 20 \times e^{\left( \frac{1 \times \Delta G_{vp6,T}^0 + 24 \times \Delta G_{vp6,PQRS}^0}{RT} \right)} [vp6_{trimers}]^{13}$

**n** – represents the species involved in the assembly of a vp6 structural subunit; **Build-up** – number of ways of forming a species  $n$  from the preceding stable intermediate,  $n-1$ , and free subunit  $n=1$ ; **Build-down** – number of ways of dissociating the intermediate species  $n$  into equivalent reactants,  $n-1$ ;  $S_{1,n}$  – ratio Build-up/Build-down;  $S_{2,n} = 1$  – degeneracy of the incoming subunit (each subunit consists of 1 icosahedral asymmetric unit);  $N_{c,n}$  – number of contacts a vp6 trimer does with adjacent trimers and the vp2 layer at the  $n$ -th species;  $\Delta G_{vp6,T}^0$  – Gibbs free energy associated with the contact between vp6 trimer  $T$  and the icosahedral vp2 layer (cal.mol<sup>-1</sup>);  $\Delta G_{vp6,PQRS}^0$  – Gibbs free energy associated with the contact between vp6 trimers (cal.mol<sup>-1</sup>); **R** – gas constant (1.987 cal.mol<sup>-1</sup>K<sup>-1</sup>); **T** – temperature (K); **[n]** – concentration of each intermediate species  $n$  (M); **[vp6<sub>trimers</sub>]** – concentration of vp6 trimers, the building blocks of vp6 structural subunit (M).

Table 11b. The assembly of vp6 subunits on top of SLPs.

n	Model	Build up down	S <sub>LP</sub>	N <sub>LP</sub>	[n]
20					SLPs
21		20	1	20/1	1
22		3	1	3/1	2
23		4	2	4/2	2
24		2	2	2/2	2
25		1	5	1/5	3
26		5	1	5/1	2
27		2	2	2/2	2
28		1	2	1/2	3
29		4	1	4/1	2
30		1	6	1/6	3
					$[20] \times 13^{10} \times 240 \times e^{\left(-\frac{22 \times \Delta G_{DLP}^0}{RT}\right)} [free\ vp6]^{10}$
31		6	1	6/1	2
32		1	4	1/4	3
33		2	1	2/1	2
34		2	2	2/2	3
35		1	5	1/5	3
36		5	1	5/1	2
37		2	2	2/2	3
38		2	4	2/4	3
39		2	3	2/3	3
40		1	20	1/20	4
					$[20] \times 13^{20} \times 2 \times e^{\left(-\frac{50 \times \Delta G_{DLP}^0}{RT}\right)} [free\ vp6]^{20}$

n – represents the species involved in the assembly of vp6 layer; **Build-up** – number of ways of forming a species n from the preceding stable intermediate, n-1, and free subunit n=1; **Build-down** – number of ways of dissociating the intermediate species n into equivalent reactants, n-1; S<sub>LP</sub> – ratio Build-up/Build-down; N<sub>LP</sub> = 13 – degeneracy of the incoming subunit (each subunit consists of 13 icosahedral asymmetric unit); N<sub>VP6</sub> – number of contacts a vp6 structural subunit does with adjacent subunits and the vp2 layer at the n-th species;  $\Delta G_{VP6}^0 = \Delta G_{VP6}^0 + 24 \times \Delta G_{VP6}^0$  – Gibbs free energy of vp6 subunit association (cal.mol<sup>-1</sup>); R – gas constant (1.987 cal.mol<sup>-1</sup>.K<sup>-1</sup>); T – temperature (K); [n] – concentration of each intermediate species n (M); [free vp6] = [vp6]<sub>dimers</sub><sup>1/2</sup> – concentration of unassociated vp6 subunits at equilibrium (M).



Table 11b lists the assembly intermediates and factors describing the assembly of the second RLP layer. Modeling parameters for this layer are:  $n=20$  - vp6 layer has icosahedral geometry with 20 structural subunits;  $S_{2,n}=13$  - there are 13 equivalent orientations for the each trimer in the incoming subunit,  $z=3$  - each asymmetric unit is formed by 3 proteins;  $\Delta G^0_{DLP} = -4.08 \text{ kcal.mol}^{-1}$  [36, 37].

### 3.2.3. Formation of the vp7 layer

The structure of the outer RLP layer is similar to the vp6 layer: 780 molecules of vp7 arranged as 260 trimers in a  $T=13$  quasi-equivalent icosahedral geometry. This layer has 20 stable subunits, each one formed by 13 vp7 trimers. The assembly of vp7 trimers into a structural subunit and this into a complete layer obeys to a specific sequence similar to the assembly of the vp6 layer.

In the native virus, there are solvent channels that go across the rotavirus structure, through which mRNA transcripts emerge from the viral particle and aqueous material and biochemical substrates are transported into and out of the capsid; these channels persist even after the assembly of the outer layer. This suggest that vp7 trimers interact in a one-to-one fashion with the vp6 trimers in the inner layer by contacting top of one with bottom of the other. Here, the same type of interaction was assumed; vp7 trimer  $T$  contacts the outer DLP layer at the 3-fold axes at vp6 trimer  $T$  locus. This contact is the strongest ( $\Delta G^0_{vp7,T}$ ) whereas the vp7 trimer-trimer interactions ( $\Delta G^0_{vp7,PQRS}$ ) and the interaction between vp7 trimers P of adjacent subunits ( $\Delta G^0_{RLP}$ ) are weaker.

The solubilisation of the outer layer from the RLP is achieved using low calcium concentration ( $\sim 1 \text{ mM}$ ), by the addition of chelating agents such as EDTA or EGTA at concentrations around 1-10 mM or by lowering the

temperature [4]. No significant effect is observed on SLPs or DLPs. These results suggest that

$$\Delta G^0_{\text{SLP}} \cong \Delta G^0_{\text{DLP}} > \Delta G^0_{\text{RLP}}$$

Table 12 lists the assembly intermediates and factors describing the assembly of vp7 RLP layer.






















In virtue of model simplification and computation time, the assembly of vp7 structural subunits is lumped; hence, the Gibbs free energy of vp7 subunit association can be defined as

$$\Delta G^0_{\text{RLP}} = \Delta G^0_{\text{vp7,T}} + 24 \times \Delta G^0_{\text{vp7,PQRS}}$$

and **[free vp7]** is the concentration of unassociated vp7 subunits at equilibrium (M).

Modeling parameters for this layer are: **n=20** - vp7 layer has icosahedral geometry with 20 structural subunits; **S<sub>2,n</sub>=13** - there are 13 equivalent orientations for the each trimer in the incoming subunit, **z=3** - each asymmetric unit is formed by 3 proteins;  $\Delta G^0_{\text{RLP}} = -4.08 \text{ kcal.mol}^{-1}$  [36, 37].

**Table 12.** The assembly of vp7 subunits on top of DLPs.

n	Model	Build up	Build down	S <sub>1,n</sub>	N <sub>ca,n</sub>	[n]
40						DLPs
41		20	1	20/1	1	$[40] \times 13 \times 20 \times e^{\left(-\frac{1 \times \Delta G_{DLP}^0}{RT}\right)} [free\ vp7]^1$
42		3	1	3/1	2	$[40] \times 13^2 \times 60 \times e^{\left(-\frac{3 \times \Delta G_{DLP}^0}{RT}\right)} [free\ vp7]^2$
43		4	2	4/2	2	$[40] \times 13^3 \times 120 \times e^{\left(-\frac{5 \times \Delta G_{DLP}^0}{RT}\right)} [free\ vp7]^3$
44		2	2	2/2	2	$[40] \times 13^4 \times 120 \times e^{\left(-\frac{7 \times \Delta G_{DLP}^0}{RT}\right)} [free\ vp7]^4$
45		1	5	1/5	3	$[40] \times 13^5 \times 24 \times e^{\left(-\frac{10 \times \Delta G_{DLP}^0}{RT}\right)} [free\ vp7]^5$
46		5	1	5/1	2	$[40] \times 13^6 \times 120 \times e^{\left(-\frac{12 \times \Delta G_{DLP}^0}{RT}\right)} [free\ vp7]^6$
47		2	2	2/2	2	$[40] \times 13^7 \times 120 \times e^{\left(-\frac{14 \times \Delta G_{DLP}^0}{RT}\right)} [free\ vp7]^7$
48		1	2	1/2	3	$[40] \times 13^8 \times 60 \times e^{\left(-\frac{17 \times \Delta G_{DLP}^0}{RT}\right)} [free\ vp7]^8$
49		4	1	4/1	2	$[40] \times 13^9 \times 240 \times e^{\left(-\frac{19 \times \Delta G_{DLP}^0}{RT}\right)} [free\ vp7]^9$
50		1	6	1/6	3	$[40] \times 13^{10} \times 40 \times e^{\left(-\frac{22 \times \Delta G_{DLP}^0}{RT}\right)} [free\ vp7]^{10}$
51		6	1	6/1	2	$[40] \times 13^{11} \times 240 \times e^{\left(-\frac{24 \times \Delta G_{DLP}^0}{RT}\right)} [free\ vp7]^{11}$
52		1	4	1/4	3	$[40] \times 13^{12} \times 60 \times e^{\left(-\frac{27 \times \Delta G_{DLP}^0}{RT}\right)} [free\ vp7]^{12}$
53		2	1	2/1	2	$[40] \times 13^{13} \times 120 \times e^{\left(-\frac{29 \times \Delta G_{DLP}^0}{RT}\right)} [free\ vp7]^{13}$
54		2	2	2/2	3	$[40] \times 13^{14} \times 120 \times e^{\left(-\frac{32 \times \Delta G_{DLP}^0}{RT}\right)} [free\ vp7]^{14}$
55		1	5	1/5	3	$[40] \times 13^{15} \times 24 \times e^{\left(-\frac{35 \times \Delta G_{DLP}^0}{RT}\right)} [free\ vp7]^{15}$
56		5	1	5/1	2	$[40] \times 13^{16} \times 120 \times e^{\left(-\frac{37 \times \Delta G_{DLP}^0}{RT}\right)} [free\ vp7]^{16}$
57		2	2	2/2	3	$[40] \times 13^{17} \times 120 \times e^{\left(-\frac{40 \times \Delta G_{DLP}^0}{RT}\right)} [free\ vp7]^{17}$
58		2	4	2/4	3	$[40] \times 13^{18} \times 60 \times e^{\left(-\frac{43 \times \Delta G_{DLP}^0}{RT}\right)} [free\ vp7]^{18}$
59		2	3	2/3	3	$[40] \times 13^{19} \times 40 \times e^{\left(-\frac{46 \times \Delta G_{DLP}^0}{RT}\right)} [free\ vp7]^{19}$
60		1	20	1/20	4	$[40] \times 13^{20} \times 2 \times e^{\left(-\frac{50 \times \Delta G_{DLP}^0}{RT}\right)} [free\ vp7]^{20}$

n – represents the species involved in the assembly of vp7 layer, **Build-up** – number of ways of forming a species n from the preceding stable intermediate, n-1, and free subunit n-1; **Build-down** – number of ways of dissociating the intermediate species n into equivalent reactants, n-1; S<sub>1,n</sub> – ratio Build-up/Build-down; S<sub>2,n</sub> = 13 – degeneracy of the incoming subunit (each subunit consists of 13 icosahedral asymmetric unit); N<sub>ca,n</sub> – number of contacts a vp7 structural subunit does with adjacent subunits and the vp6 layer of DLPs at the n-th species;  $\Delta G_{DLP}^0 = \Delta G_{vp7}^0 + 24 \times \Delta G_{vp7}^{pos}$  – Gibbs free energy of vp7 subunit association (cal.mol<sup>-1</sup>); R – gas constant (1.987 cal.mol<sup>-1</sup>.K<sup>-1</sup>); T – temperature (K); [n] – concentration of each intermediate species n (M); [free vp7] – concentration of unassociated vp7 subunits at equilibrium (M).

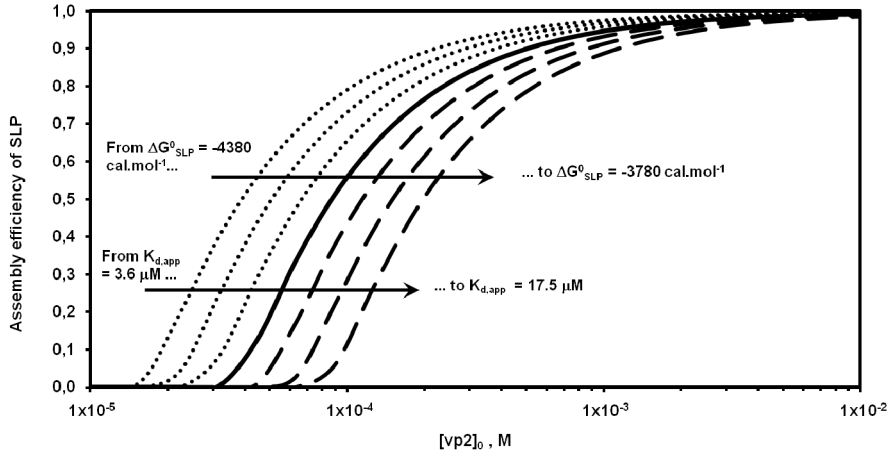
## 4. RESULTS AND DISCUSSION

### 4.1. Pseudo-critical protein concentration

Model simulations show that initial concentrations of protein structural subunits play a critical role on particle assembly. A small change in their concentrations has a profound effect on assembly efficiency. Likewise, a small change in  $\Delta G_n^0$  may result in a significant degradation or improvement in particle assembly. Such variations can be illustrated by analysing the simplest case of SLP formation. Figure 23 shows the SLP assembly efficiency as function of the initial vp2 concentration,  $[\text{vp2}]_0$ , for different values of Gibbs free energy of vp2 subunit association ( $\Delta G_{\text{SLP}}^0$ ). The assembly efficiency curves follow a sigmoidal shape with respect to  $[\text{vp2}]_0$ . Independently of  $\Delta G_{\text{SLP}}^0$ , few particles are formed at very low  $[\text{vp2}]_0$ ; above a given critical  $[\text{vp2}]_0$  value, most subunits assemble into macrostructures. This threshold represents the minimum initial concentration of vp2 needed to ensure a measurable SLP formation, i.e., a pseudo-critical concentration,  $K_{\text{d,app}}$ , that satisfies the constraint  $[1] = [20] = K_{\text{d,app}}$  at equilibrium for a specific  $\Delta G_{\text{SLP}}^0$ ;  $K_{\text{d,app}}$  can be derived from Table 10:

$$K_{\text{d,app}} = \exp \left( \frac{30\Delta G_{\text{SLP}}^0 - RT \ln(3^{19}/20)}{19RT} \right) \quad \text{Eq. 4.7}$$

It is clear that  $K_{\text{d,app}}$  increases with the increase in  $\Delta G_{\text{SLP}}^0$ ;  $K_{\text{d,app}} = 3.6 \mu\text{M}$  for  $\Delta G_{\text{SLP}}^0 = -4.38 \text{ kcal.mol}^{-1}$  whereas at a higher  $\Delta G_{\text{SLP}}^0$ ,  $-3.78 \text{ kcal.mol}^{-1}$ , it increases almost 4-fold to  $17.5 \mu\text{M}$ . These results confirm that there is a crisp transition between unassembled structural subunits and intact icosahedra.



**Figure 23.** SLP assembly efficiency as function of the initial vp2 concentration,  $[vp2]_0$ , for different Gibbs free energy values,  $\Delta G^0_{SLP}$ . The full line represents  $\Delta G^0_{SLP} = -4.08 \text{ kcal.mol}^{-1}$  as estimated by Zlotnick 1994, the dash lines represent  $\Delta G^0_{SLP} > -4.08 \text{ kcal.mol}^{-1}$  and the dot lines represent  $\Delta G^0_{SLP} < -4.08 \text{ kcal.mol}^{-1}$ .  $K_{d,app}$  represents the pseudo-critical concentration that satisfies the constraint **[1]** (concentration of unassociated vp2 subunits at equilibrium -  $[1] = [svp2]_e = [20]$  (complete SLP) =  $K_{d,app}$  at equilibrium for a specific  $\Delta G^0_{SLP}$ ).

#### 4.2. The effect of the number of association contacts

The aforementioned pseudo-critical concentration  $K_{d,app}$  can be also interpreted as the dissociation constant for a single contact in the final structure when the concentration of species **1** and **20** coincide **[37]**. As shown in Figure 23, decreasing  $\Delta G^0_{SLP}$  also decreases the  $K_{d,app}$ , suggesting that the reaction  $SLP \rightarrow 20 \times vp2$  is not favoured. In other words, the more negative is  $\Delta G^0_{SLP}$  the higher is the association constant for a subunit-subunit interaction,  $K_{contact}$  **[41]**:

$$\Delta G^0_{SLP} = -RT \times \ln(K_{contact}) \quad \text{Eq. 4.8}$$

Since the overall SLP assembly constant,  $K_{layer}$ , is proportional to  $K_{contact}$  **[41]**:

$$K_{\text{layer}} \approx K_{\text{contact}}^{\sum_{n=1}^{20} N_{c,n}} \quad \text{Eq.4.9}$$

an increase in  $K_{\text{contact}}$  favours the reaction  $20 \times \text{vp2} \rightarrow \text{SLP}$ , increasing species **20** formation and consequently the assembly efficiency. In the end, equations 4.8 and 4.9 clearly demonstrate that as more contacts are established *per* subunit, the lower is  $K_{d,\text{app}}$  i.e., more stable is the particle formed.

### 4.3. Simulation of RLP assembly

The concentration of species **1** to **60** at equilibrium can be calculated as functions of initial protein concentrations in solution before equilibrium. For given initial protein concentrations  $[\text{vp2}]_0$ ,  $[\text{vp6}]_0$  and  $[\text{vp7}]_0$ , the concentrations of intermediate species needs to be calculated iteratively to ensure that protein material balances are obeyed, i.e. the initial amount of vp2, vp6 and vp7 must be equal to the sum of the amounts of proteins incorporated in all intermediate species. This translates into the following balance equations:

$$[\text{vp2}]_0 = [\text{svp2}]_e S_{2,n} z + \sum_{n=2}^{20} [n] m_n S_{2,n} z + \sum_{n=21}^{60} [n] 20 S_{2,n} z \quad \text{Eq. 4.10}$$

$$[\text{vp6}]_0 = [\text{svp6}]_e S_{2,n} z + \sum_{n=21}^{40} [n] (m_n - 20) S_{2,n} z + \sum_{n=41}^{60} [n] 20 S_{2,n} z \quad \text{Eq. 4.11}$$

$$[\text{vp7}]_0 = [\text{svp7}]_e S_{2,n} z + \sum_{n=41}^{60} [n] (m_n - 40) S_{2,n} z \quad \text{Eq. 4.12}$$

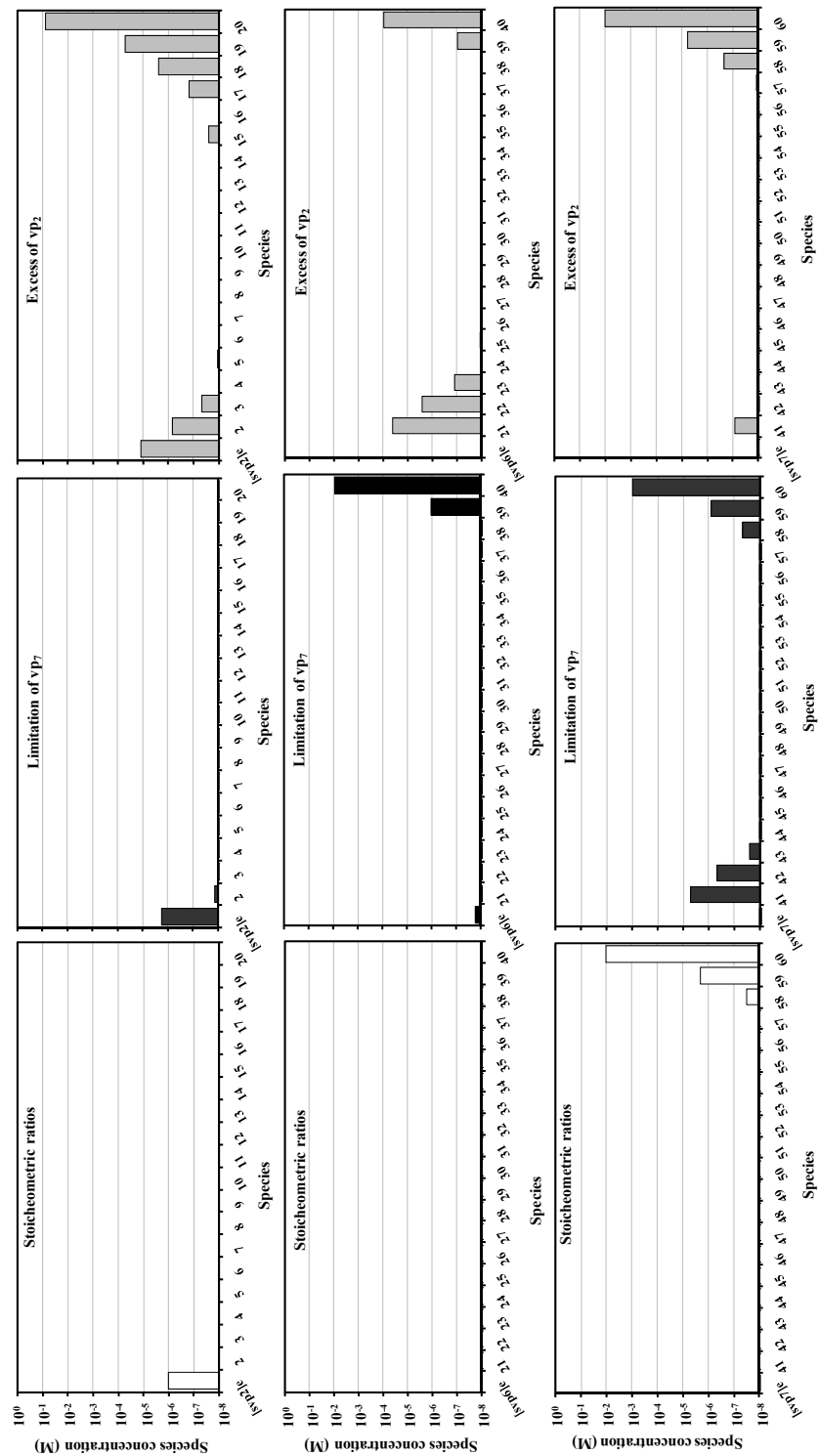
where  $[\text{svp2}]_e$ ,  $[\text{svp6}]_e$  and  $[\text{svp7}]_e$  are the concentrations of unassociated vp2, vp6 and vp7 subunits at equilibrium, respectively,  $[n]$  the concentration of species  $n$ ,  $m_n$  the number of structural subunits in species  $n$ ,  $S_{2,n}$  the degeneracy of the incoming subunit in species  $n$  and  $z$  the number of proteins in each oligomer. Parameters  $S_{2,n}$  and  $z$  are specific for each layer:  $S_{2,n} = 3$  and  $z = 2$  (vp2 layer);  $S_{2,n} = 13$  and  $z = 3$  (vp6 and vp7 layers) (see Section 3 for details). This system can be solved easily by manipulating

iteratively the free amounts of vp2, vp6 and vp7 until the balance equations 4.10 to 4.12 are obeyed to a given tolerance.

RLP assembly simulations were performed for three initial protein concentration scenarios: 1)  $[\text{vp2}]_0 = 1.2 \text{ M}$ ,  $[\text{vp6}]_0 = 7.8 \text{ M}$  and  $[\text{vp7}]_0 = 7.8 \text{ M}$ ; 2)  $[\text{vp2}]_0 = 1.2 \text{ M}$ ,  $[\text{vp6}]_0 = 7.8 \text{ M}$  and  $[\text{vp7}]_0 = 0.78 \text{ M}$ ; 3)  $[\text{vp2}]_0 = 12 \text{ M}$ ,  $[\text{vp6}]_0 = 7.8 \text{ M}$  and  $[\text{vp7}]_0 = 7.8 \text{ M}$ . The rationale for this choice of concentrations is testing either protein limitation or excess in relation to the RLP stoichiometric composition (vp2:vp6:vp7 - 120:780:780 molecules). In this way, vp6 is provided at the exact stoichiometric ratio while scenarios of vp7 limitation and excess of vp2 are also explored. The Gibbs free energy of subunit association was assumed to be equal for all layers ( $\Delta G^0_{\text{SLP}} = \Delta G^0_{\text{DLP}} = \Delta G^0_{\text{RLP}} = -4.08 \text{ kcal.mol}^{-1}$ ). Figure 24 shows the particles distribution at equilibrium for the three initial protein concentration scenarios allowing to make the following observations:

- When initial protein concentrations (vp2, vp6 and vp7) obey the RLP stoichiometric ratio, the distribution of particles at equilibrium is dominated by a very small number of more stable structures, namely subunit **1** (i.e.  $[\text{1}] = [\text{svp2}]_0$ ) and structure **60** (complete RLP). The remaining intermediates are present but at residual concentrations. This shows that, for low Gibbs free energies, thermodynamics favours the formation of intact icosahedra whenever the respective proteins are not limiting.

- When vp7 is limiting, and according to Le Chatelier's principle, it should be expected that the equilibrium composition of the lumped reaction  $\text{DLP} + 20 \times \text{vp7} \rightleftharpoons \text{RLP}$  shifts to the left with more DLP and less RLP formed. Indeed, when comparing the concentration of structure **40** (DLP) in this case with the one where proteins are provided at the exact stoichiometric ratios, a significant increase is observed; the opposite occurs to the concentration of structure **60**.



**Figure 24.** The concentration of assembly intermediates for three initial protein concentration scenarios: 1)  $[vp2]_0 = 1.2$  M,  $[vp6]_0 = 7.8$  M and  $[vp7]_0 = 7.8$  M (stoichiometric ratios – white bars); 2)  $[vp2]_0 = 1.2$  M,  $[vp6]_0 = 0.78$  M and  $[vp7]_0 = 7.8$  M (limitation of vp7 – black bars); 3)  $[vp2]_0 = 12$  M,  $[vp6]_0 = 7.8$  M and  $[vp7]_0 = 7.8$  M (excess of vp2 – grey bars). The Gibbs free energy of subunit association was assumed to be equal for all layers ( $\Delta G^0_{SLP} = \Delta G^0_{DLP} = \Delta G^0_{RLP} = -4.08$  kcal.mol $^{-1}$ ).

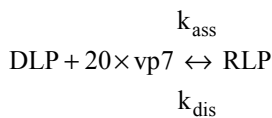


Concomitantly, moderate sized structures such as species **39**, **41** or **42** start to appear at significant concentrations. This demonstrates that protein limitation negatively impacts particle assembly even beyond the stoichiometric rational, since intermediate structures accumulate which represent unnecessary waste of proteins.

- An excess of vp2 concentration shifts the equilibrium of the lumped reaction  $20 \times \text{vp2} \rightleftharpoons \text{SLP}$  towards the generation of more SLP (Le Chatelier's principle). For this reason, the concentration of structure **20** (SLP) is higher than the one obtained when proteins are provided at the exact stoichiometric ratios. This equilibrium shift has also a profound impact on moderate sized structures such as species **17** to **23** or **39** to **41**; their concentrations increase significantly. The high concentrations observed for intermediate structures **20**, **17-23** and **39-41** indicate that the reaction  $20 \times \text{vp2} \rightleftharpoons \text{SLP} + 20 \times \text{vp6} \rightleftharpoons \text{DLP} + 20 \times \text{vp7} \rightleftharpoons \text{RLP}$  is not product-favoured. This demonstrates that proteins in excess impact negatively on final RLP yields while reducing the assembly efficiency due to the accumulation of unassembled proteins.

#### 4.4. *In vitro* RLPs assembly experiments

The *in vitro* assembly of RLPs can be described by the lumped reaction



where  $k_{\text{ass}}$  and  $k_{\text{dis}}$  are the reaction rate constants of RLP assembly and disassembly, respectively. The ratio of the reaction rate constants is known as the equilibrium constant,  $K_{\text{eq}}$ :

$$K_{\text{eq}} = \frac{k_{\text{ass}}}{k_{\text{dis}}} = \frac{[\text{RLP}]}{[\text{DLP}] \times [\text{vp7}]^{20}} \quad \text{Eq. 4.13}$$

Based on equation 4.9, the equilibrium constant  $K_{\text{eq}}$  is related with the

Gibbs free energy of vp7 subunit association,  $\Delta G_{\text{RLP}}^0$ , by:

$$\Delta G_{\text{RLP}}^0 = -RT \times \ln(K_{\text{eq}}) \times \left(1 / \sum_{n=40}^{60} N_{c,n}\right) \quad \text{Eq. 4.14}$$

The Gibbs free energy of vp7 subunit association for the *in vitro* RLPs assembly experiments was calculated using the formula (see Table 12):

$$\Delta G_{\text{RLP}}^0 = \ln \left( \frac{[40] \times 13^{20} \times 2 \times [\text{free vp7}]^{20}}{[60]} \right) \times \frac{RT}{50} \quad \text{Eq. 4.15}$$

where [40] is the concentration of DLP at equilibrium,  $[\text{DLP}]_e$ , [free vp7] is the concentration of unassociated vp7 subunits at equilibrium,  $[\text{svp7}]_e$  and [60] is the concentration of RLP at equilibrium,  $[\text{RLP}]_e$ . The obtained results are summarized in Table 13.

Results show that the increase in pH, [NaCl] or  $[\text{Ca}^{2+}]$  had a negative impact on the Gibbs free energy of vp7 subunit association; the  $\Delta G_{\text{RLP}}^0$  values are higher than the one estimated for the standard condition of assembly (pH 5.5, 25 °C, 0.1 M of NaCl and 1 mM of  $\text{Ca}^{2+}$ ). This indicates that the equilibrium constants for the assembly reactions at pH 8, [NaCl] = 0.5 M or  $[\text{Ca}^{2+}] = 5$  mM decreased and, if  $K_{\text{eq}}$  decreased, product formation became less favoured; hence, the assembly efficiencies are lower. The effect of temperature variation on Gibbs free energy of vp7 subunit association and assembly efficiency is discussed below.

#### 4.5. Temperature dependency

The dependency between temperature **T** and the equilibrium constant  $K_{\text{eq}}$  is described by the van't Hoff equation:

$$\frac{\partial \ln(K_{\text{eq}})}{\partial (1/T)} = -\frac{\Delta H^0}{R} \quad \text{Eq. 4.16}$$

**Table 13.** *In vitro* assembly and disassembly of RLPs.

Assembly	Assembly condition		Ass <sub>eff</sub>	Temperature (K)	ΔG <sup>0</sup> <sub>RLP</sub> (cal.mol <sup>-1</sup> )	ΔG <sup>0</sup> <sub>layer, vp7</sub> (kcal.mol <sup>-1</sup> )
	pH 5.5 / [NaCl]=0.1 M / [Ca <sup>2+</sup> ]=1 mM		0.48	298.15	-4830	-242
	pH 8.0 / [NaCl]=0.1 M / [Ca <sup>2+</sup> ]=1 mM					
	pH 5.5 / [NaCl]=0.5 M / [Ca <sup>2+</sup> ]=1 mM		< 0.10	298.15	> -4675	> -234
	pH 5.5 / [NaCl]=0.1 M / [Ca <sup>2+</sup> ]=5 mM					
	pH 5.5 / [NaCl]=0.1 M / [Ca <sup>2+</sup> ]=1 mM		< 0.10	308.15	> -4832	> -242
Disassembly	Disassembly condition		Dis <sub>eff</sub>	Temperature (K)	ΔG <sup>0</sup> <sub>RLP</sub> (cal.mol <sup>-1</sup> )	ΔG <sup>0</sup> <sub>layer, vp7</sub> (kcal.mol <sup>-1</sup> )
	TNC buffer	EGTA 1 mM	0.72	308.15	-3298	-165
		EDTA 1 mM	0.72		-3297	-165
		EGTA 2 mM	0.22		-3609	-180
		EDTA 2 mM	0.19		-3655	-183
	D-PBS buffer	EGTA 1 mM	0.66	308.15	-3320	-166
		EDTA 1 mM	0.46		-3420	-171
		EGTA 2 mM	0.33		-3507	-175
		EDTA 2 mM	0.99		-3175	-159
	Storage time	4 months	0.75	277.15	-2947	-147
		12 months	0.81		-2928	-146
15 months		0.85	-2915		-146	
Reassembly	Disassembly condition	Reassembly condition	Ass <sub>eff</sub>	Temperature (K)	ΔG <sup>0</sup> <sub>RLP</sub> (cal.mol <sup>-1</sup> )	ΔG <sup>0</sup> <sub>layer, vp7</sub> (kcal.mol <sup>-1</sup> )
	EGTA 2 mM (TNC)		0.47	298.15	-3626	-181
	EDTA 2 mM (TNC)	[Ca <sup>2+</sup> ]=5 mM and T=25 °C	0.28		-3586	-179
	EGTA 2 mM (D-PBS)		0.54		-3570	-179
	EDTA 2 mM (D-PBS)		0.80		-3524	-176

Ass<sub>eff</sub> - assembly efficiency; [DLP]<sub>0</sub> - initial concentration of DLP (M); [DLP]<sub>e</sub> - concentration of DLP at equilibrium (M); Dis<sub>eff</sub> - disassembly efficiency; [RLP]<sub>0</sub> - initial concentration of RLP (M); [RLP]<sub>e</sub> - concentration of RLP at equilibrium (M);  $\Delta G^0_{RLP}$  - Gibbs free energy of vp7 subunit association (cal.mol<sup>-1</sup>);  $\Delta G^0_{layer, vp7}$  - Gibbs free energy of vp7 layer assembly on top of DLPs; [NaCl] - sodium chloride concentration (M); [Ca<sup>2+</sup>] - calcium concentration (M); T - temperature (K)

where  $\Delta H^0$  is the standard enthalpy of reaction ( $\text{cal.mol}^{-1}$ ). The  $\Delta H^0$  and the standard entropy of reaction,  $\Delta S^0$  ( $\text{cal.mol}^{-1}.\text{K}^{-1}$ ), can be calculated by plotting the natural logarithm of the equilibrium constant versus the reciprocal temperature as follows:

$$\Delta G^0 = \Delta H^0 - T\Delta S^0 \quad \text{Eq. 4.17}$$

$$\Delta G^0 = -RT \times \ln(K_{\text{eq}}) \quad \text{Eq. 4.18}$$

$$\ln(K_{\text{eq}}) = \left( -\frac{\Delta H^0}{R} \right) \times \frac{1}{T} + \frac{\Delta S^0}{R} \quad \text{Eq. 4.19}$$

where the slope is equal to minus the standard enthalpy change divided by the gas constant and the intercept is equal to the standard entropy change divided by the gas constant.

Based on equation 19 and the *in vitro* assembly experiments carried out at 25 °C and 35 °C (Table 13), the standard entropy and enthalpy of vp7 layer assembly on top of DLPs,  $\Delta S^0_{\text{layer, vp7}}$  and  $\Delta H^0_{\text{layer, vp7}}$  respectively, were calculated. The  $\Delta S^0_{\text{layer, vp7}}$  and  $\Delta H^0_{\text{layer, vp7}}$  of the reaction **DLP + 20 × vp7**  $\rightleftharpoons$  **RLP** are positive ( $\sim 100 \text{ cal.mol}^{-1}.\text{K}^{-1}$ ) and negative ( $\sim -240 \text{ kcal.mol}^{-1}$ ) respectively, meaning that the reaction is spontaneous and exothermic, and consistent with  $\Delta G^0_{\text{RLP}} \sim -5 \text{ kcal.mol}^{-1}$ . The entropy contribution to the Gibbs free energy of vp7 layer assembly on top of DLPs,  $\Delta G^0_{\text{layer, vp7}}$ , is smaller than that of enthalpy, which is reflected in the fact that while the association changes significantly with the temperature,  $\Delta G^0_{\text{layer, vp7}}$  does not, i.e., the association is enthalpy driven. When  $\Delta H^0 < 0$ , then  $K_{\text{eq}}$  decreases with increasing temperature, in accordance with Le Chatelier's principle. If  $K_{\text{eq}}$  decreases, RLP formation is less favoured and, consequently, the assembly efficiency decreases. For this reason, the increase in temperature reduced significantly the efficiency of RLP assembly (Table 13).

#### 4.6. Interpretation of *in vitro* disassembly data

*In vitro* disassembly experiments were performed at different concentrations of EGTA and EDTA (see Table 13). For each experiment, the Gibbs free energy of vp7 subunit association was calculated based on equation 4.15; the obtained results are summarized in Table 13.

It is clear from equations 4.13 and 4.14 that the equilibrium constant  $K_{eq}$  is inversely proportional to the Gibbs free energy of vp7 subunit association and to the disassembly efficiency; the lower the  $K_{eq}$ , the higher is the  $\Delta G^0_{RLP}$  and the disassembly efficiency. Therefore, the optimal condition to disassemble RLP is to use EDTA 2 mM in D-PBS as it presents the highest  $\Delta G^0_{RLP}$  (Table 13). Interestingly, the concentration of chelating agent appears to be correlated with the  $\Delta G^0_{RLP}$ ; higher concentrations induce lower  $\Delta G^0_{RLP}$ . The exception is the case of EDTA 2 mM in D-PBS where  $\Delta G^0_{RLP}$  is higher than the one achieved at EDTA 1 mM.

Upon the disassembly of RLPs with EDTA and EGTA 2 mM, reassembly was performed with dialysis against buffer with calcium 5 mM and decreasing the temperature to 25 °C (Table 13). As a rule, the  $\Delta G^0_{RLP}$  values for the reassembly experiments are lower than the ones obtained for the disassembly. This happens because lowering the temperature in an exothermic reaction enhances product formation by shifting the equilibrium to the right ( $DLP + 20 \times vp7 \rightleftharpoons RLP$ ). Moreover, by increasing the concentration of calcium, the vp7 attachment to DLP is increased. However, very high concentrations of calcium (5 mM; see Table 13) seem to have a deleterious effect on the association. Summing up, the combined effect of optimal calcium concentration and temperature resulted in the increase in  $K_{eq}$  and the subsequent decrease in  $\Delta G^0_{RLP}$ .

The effect of storage time on  $\Delta G^0_{\text{RLP}}$  was also evaluated; results are summarized in Table 13. The  $\Delta G^0_{\text{RLP}}$  increases with storage time, indicating that the transition from RLP to DLP is favoured as time progresses ( $K_{\text{eq}}$  decreases). The highest disassembly efficiency is achieved at a storage time of 15 months which corresponds to the highest  $\Delta G^0_{\text{RLP}}$ .

#### 4.7. Design of optimal thermodynamic conditions for global RLP assembly

The assembly of protein complexes such as RLP can be optimized through the manipulation of viral protein concentrations and solution properties with impact on Gibbs free energy of subunit association (such as temperature, pH, ionic strength and particular ions concentration, namely  $\text{Ca}^{2+}$ ). In order to find the theoretical optimal conditions for RLP assembly, model simulations were performed for different values of Gibbs free energy of subunit association and vp6 and vp7 initial protein concentrations,  $[\text{vp6}]_0$  and  $[\text{vp7}]_0$ , respectively. The concentration of vp2,  $[\text{vp2}]_0$ , was kept constant at 1 M and the Gibbs free energies of vp2, vp6 and vp7 subunit association were considered equal ( $\Delta G^0_{\text{SLP}} = \Delta G^0_{\text{DLP}} = \Delta G^0_{\text{RLP}}$ ).

The RLP assembly efficiency is defined as the ratio between the mass of intact RLP and the mass of initial proteins. To assess the implications of equilibrium, the assembly efficiency predicted by the thermodynamic model was compared to that of an irreversible assembly reaction based on the mass composition of RLP:  $1 \times \text{vp2} + 2.84 \times \text{vp6} + 2.35 \times \text{vp7} \rightarrow 6.19 \times \text{RLP}$ .

In the irreversible assembly model, the efficiency of assembly will be controlled by the limiting protein.

Figure 25A shows the RLP assembly efficiency for the irreversible model while Figures 25B to 25F show the assembly efficiency of the

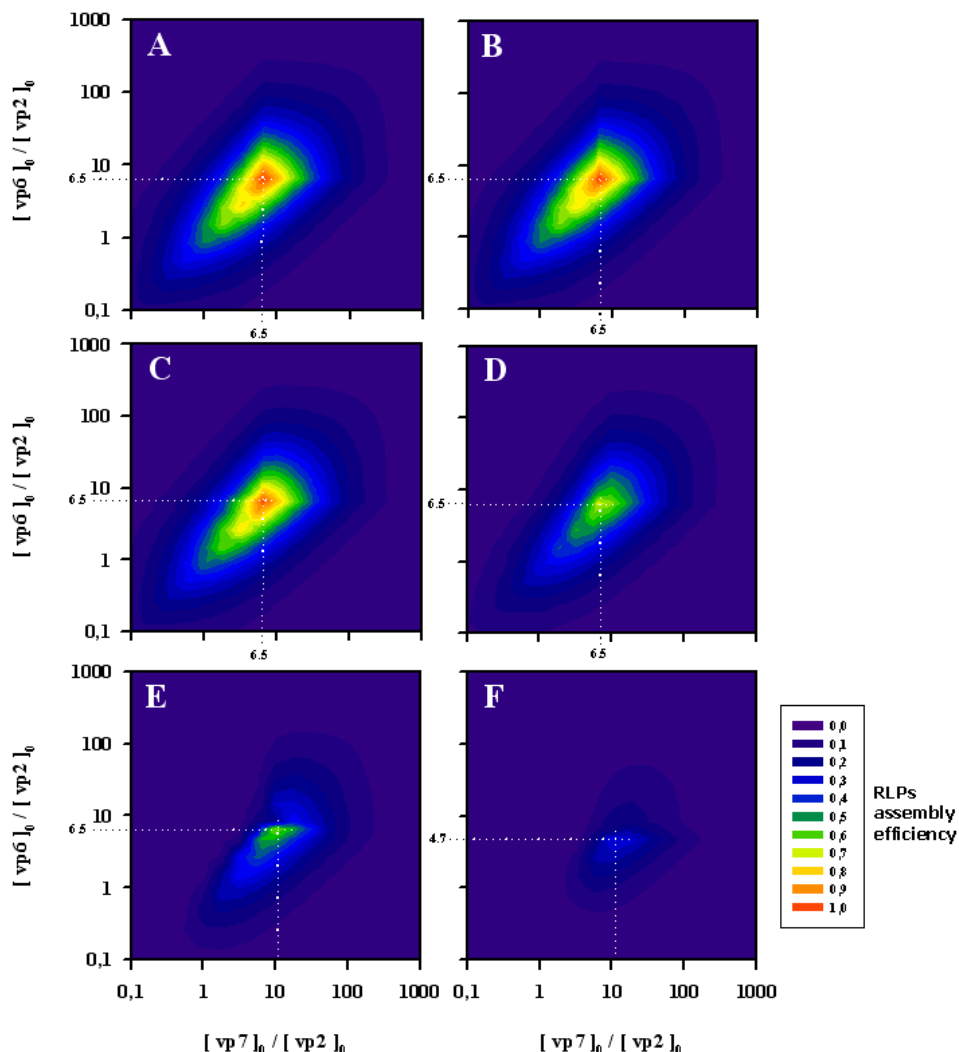
thermodynamic model for  $\Delta G^0_{\text{RLP}}$  values of -4.08, -2.58, -1.5, -0.75 and -0.408 kcal.mol<sup>-1</sup>, respectively.

As expected, the irreversible model predicts higher efficiencies than the equilibrium model; as the  $\Delta G^0_{\text{RLP}}$  decreases, the equilibrium model converges to the irreversible model.

The maximum RLP efficiency is obtained for initial proteins concentrations coincident to the particle stoichiometric ratio (vp2:vp6:vp7 - 120:780:780 molecules). This optimum is observed for the irreversible model and also for the equilibrium model with low  $\Delta G^0_{\text{RLP}}$ , namely  $\Delta G^0_{\text{RLP}} = -4.08, -2.58$  and  $-1.5$  kcal.mol<sup>-1</sup>. Significant deviations are only obtained for high  $\Delta G^0_{\text{RLP}}$  such as -0.75 and -0.408 kcal.mol<sup>-1</sup>, where the optimal point becomes  $[\text{vp6}]_0/[\text{vp2}]_0=6.5 + [\text{vp7}]_0/[\text{vp2}]_0=10.1$  and  $[\text{vp6}]_0/[\text{vp2}]_0=4.7 + [\text{vp7}]_0/[\text{vp2}]_0=10.1$ , respectively.

This shows that, for low Gibbs free energies, thermodynamics maximizes the formation of intact icosahedra whenever vp2, vp6 and vp7 are provided at the exact RLP stoichiometric ratio. At high Gibbs free energies, there is a translocation of proteins pseudo-critical concentrations; so, optimum concentrations are no longer coincident to the stoichiometric ratio of RLPs.

A noteworthy result is that the dependency between assembly efficiency and Gibbs free energy of subunit association is not linear. For values of Gibbs free energy of subunit association below -2 kcal.mol<sup>-1</sup>, changes in Gibbs free energy have no significant effect on the assembly efficiency since the initial concentrations of proteins are above a pseudo-critical value, which guarantees that particle formation is maximized. Above this threshold, the limiting factor becomes the Gibbs free energy of subunit association; hence, as Gibbs free energy increases, a crisp transition from intact icosahedra to unassembled subunits is observed.



**Figure 25.** The effect of initial protein concentrations and Gibbs free energy of subunit association on RLP assembly. The initial concentration of vp2 was kept constant at 1 M while the initial concentrations of vp6 and vp7 varied from 0.1 M to 1000 M. The Gibbs free energies of vp2, vp6 and vp7 association were considered equal. Graph **A** shows the RLP assembly efficiency for the irreversible model while graphs **B** to **F** show the assembly efficiency of the thermodynamic model for  $\Delta G^0_{\text{RLP}}$  values of -4.08, -2.58, -1.5, -0.75 and -0.408 kcal.mol<sup>-1</sup>, respectively. The RLP assembly efficiency is defined as the ratio between the mass of intact RLP and the mass of initial proteins.



## 5. CONCLUSION

In this study, a thermodynamic equilibrium-based model describes for the first time the assembly of a protein complex with three different layers. Simulations showed that whenever proteins are provided at the VLP stoichiometric composition, particles distribution is dominated by structural subunits and completely formed icosahedra; remaining intermediates are present but at residual concentrations. If such stoichiometric ratio is not obeyed, several other intermediates are formed. Variations in the Gibbs free energy of subunit association impact on optimal initial proteins concentration and on assembly efficiency. For low Gibbs free energies, thermodynamics maximizes the formation of intact icosahedra whenever proteins are provided at the exact stoichiometric ratio. At high Gibbs free energies, there is a translocation of proteins pseudo-critical concentrations; so, optimum concentrations are no longer coincident to the particle stoichiometric composition. The dependency between assembly efficiency and Gibbs free energy of subunit association is not linear. For values of Gibbs free energy of subunit association below  $-2 \text{ kcal.mol}^{-1}$ , changes in Gibbs free energy do not affect significantly the assembly efficiency. Above this threshold, a crisp transition from intact icosahedra to unassembled subunits is observed. Summarizing, the concentration of correctly formed particles is maximized whenever structural proteins are provided at the exact stoichiometric ratio and at high initial concentrations. Particle malformation increases drastically when such conditions are not met. Host cells can hardly be manipulated to operate in such a narrow region, thus the formation of malformed particles is likely to be inevitable *in vivo*.

## 6. ACKNOWLEDGMENTS

The authors are grateful to Eng. Ana Luisa Escapa (IBET) for helping with the capillary zone electrophoresis analyses. The authors wish to thank Dr. Ana Luisa Simplicio (ITQB-UNL; IBET) for thoughtful discussions and to Dr. Didier Poncet and Dr. Annie Charpilienne (CNRS-INRA) for providing the recombinant baculoviruses. Dr. Max Ciarlet (Merck & Co.) is acknowledged for the capillary zone electrophoresis protocol. The authors acknowledge the financial support received from the Portuguese Foundation for Science and Technology (SFRH/BD/21910/2005 and SFRH/BD/28323/2006) and from the European Commission (Baculogenes, LSHB-2006-037541).

## 7. REFERENCES

1. Labbe, M., *et al.*, **1991**. Expression of rotavirus VP2 produces empty corelike particles. *J Virol* 6 (65): 2946-52.
2. Prasad, B.V., *et al.*, **1988**. Three-dimensional structure of rotavirus. *J Mol Biol* 2 (199): 269-75.
3. Vieira, H., *et al.*, **2005**. Triple layered rotavirus VLP assembly: kinetics of vector replication, mRNA stability and recombinant protein production. *J Biotechnol*, 1 (120): 72-82.
4. Mellado, M.C., *et al.*, **2009**. Impact of physicochemical parameters on in vitro assembly and disassembly kinetics of recombinant triple-layered rotavirus-like particles. *Biotechnol Bioeng* 4 (104): 674-86.
5. Rapaport, D.C., Johnson, J.E., and Skolnick, J., **1999**. Supramolecular self-assembly: molecular dynamics modeling of polyhedral shell formation. *Computer Physics Communications* (121-122): 231-235.
6. Bruinsma, R.F., *et al.*, **2003**. Viral self-assembly as a thermodynamic process. *Phys Rev Lett* 24 (90): 248101.
7. Reddy, V.S., *et al.*, **1998**. Energetics of quasiequivalence: computational analysis of protein-protein interactions in icosahedral viruses. *Biophys J* 1 (74): 546-58.
8. Berger, B., *et al.*, **1994**. Local rule-based theory of virus shell assembly. *Proc Natl Acad Sci U S A* 16 (91): 7732-6.
9. Kerner, R., **2005**. A stochastic model of icosahedral capsid growth. *Computational and Mathematical Methods in Medicine* 2 (6): 95-97.
10. Andrews, J.M. and Roberts, C.J., **2007**. A Lumry-Eyring nucleated polymerization model of protein aggregation kinetics: 1. Aggregation with pre-equilibrated unfolding. *J Phys Chem B* 27 (111): 7897-913.
11. Twarock, R., **2004**. A tiling approach to virus capsid assembly explaining a structural puzzle in virology. *J Theor Biol* 4 (226): 477-82.
12. Zlotnick, A., *et al.*, **2002**. A small molecule inhibits and misdirects assembly of hepatitis B virus capsids. *J Virol* 10 (76): 4848-54.
13. Zlotnick, A., *et al.*, **1999**. A theoretical model successfully identifies features of hepatitis B virus capsid assembly. *Biochemistry* 44 (38): 14644-52.
14. Endres, D., *et al.*, **2005**. A reaction landscape identifies the intermediates critical for self-assembly of virus capsids and other polyhedral structures. *Protein Sci* 6 (14): 1518-25.
15. Endres, D. and Zlotnick, A., **2002**. Model-based analysis of assembly kinetics for virus capsids or other spherical polymers. *Biophys J* 2 (83): 1217-30.

16. Zlotnick, A., *et al.*, **2000**. Mechanism of capsid assembly for an icosahedral plant virus. *Virology* 2 (277): 450-6.
17. Mellado, M.C., *et al.*, **2008**. Purification of recombinant rotavirus VP7 glycoprotein for the study of in vitro rotavirus-like particles assembly. *J Chromatogr B Analyt Technol Biomed Life Sci* 1-2 (874): 89-94.
18. Mellado, M.C., *et al.*, **2008**. Sodium dodecyl sulfate-capillary gel electrophoresis analysis of rotavirus-like particles. *J Chromatogr A* 1 (1192): 166-72.
19. Tossier, G., *et al.*, **1992**. Expression of the major capsid protein VP6 of group C rotavirus and synthesis of chimeric single-shelled particles by using recombinant baculoviruses. *J Virol* 10 (66): 5825-31.
20. Cruz, P.E., *et al.*, **1998**. Optimization of the production of virus-like particles in insect cells. *Biotechnol Bioeng* 4 (60): 408-18.
21. Peixoto, C., *et al.*, **2007**. Downstream processing of triple layered rotavirus like particles. *J Biotechnol* 3 (127): 452-61.
22. Prevelige, P.E., Jr., *et al.*, **1990**. Conformational states of the bacteriophage P22 capsid subunit in relation to self-assembly. *Biochemistry* 23 (29): 5626-33.
23. Tuma, R., Prevelige, P.E., Jr., and Thomas, G.J., Jr., **1996**. Structural transitions in the scaffolding and coat proteins of P22 virus during assembly and disassembly. *Biochemistry* 14 (35): 4619-27.
24. Butcher, S.J., Bamford, D.H., and Fuller, S.D., **1995**. DNA packaging orders the membrane of bacteriophage PRD1. *Embo J* 24 (14): 6078-86.
25. Salunke, D.M., Caspar, D.L., and Garcea, R.L., **1986**. Self-assembly of purified polyomavirus capsid protein VP1. *Cell* 6 (46): 895-904.
26. Salunke, D.M., Caspar, D.L., and Garcea, R.L., **1989**. Polymorphism in the assembly of polyomavirus capsid protein VP1. *Biophys J* 5 (56): 887-900.
27. Bancroft, J.B., **1970**. The self-assembly of spherical plant viruses. *Adv Virus Res* (16): 99-134.
28. Sorger, P.K., Stockley, P.G., and Harrison, S.C., **1986**. Structure and assembly of turnip crinkle virus. II. Mechanism of reassembly in vitro. *J Mol Biol* 4 (191): 639-58.
29. Savithri, H.S. and Erickson, J.W., **1983**. The self-assembly of the cowpea strain of southern bean mosaic virus: formation of T = 1 and T = 3 nucleoprotein particles. *Virology* 1 (126): 328-35.
30. Stirzaker, S.C., Poncet, D., and Both, G.W., **1990**. Sequences in rotavirus glycoprotein VP7 that mediate delayed translocation and retention of the protein in the endoplasmic reticulum. *J Cell Biol* 4 (111): 1343-50.
31. Au, K.S., Mattion, N.M., and Estes, M.K., **1993**. A subviral particle binding domain on the rotavirus nonstructural glycoprotein NS28. *Virology* 2 (194): 665-73.
32. Aoki, S.T., *et al.*, **2009**. Structure of rotavirus outer-layer protein VP7 bound with a neutralizing Fab. *Science* 5933 (324): 1444-7.
33. Madore, H.P., *et al.*, **1999**. Biochemical and immunologic comparison of virus-like particles for a rotavirus subunit vaccine. *Vaccine* 19 (17): 2461-2471.
34. Charpilienne, A., *et al.*, **2002**. Identification of rotavirus VP6 residues located at the interface with VP2 that are essential for capsid assembly and transcriptase activity. *J Virol* 15 (76): 7822-31.
35. Gorziglia, M., *et al.*, **1985**. Biochemical evidence for the oligomeric (possibly trimeric) structure of the major inner capsid polypeptide (45K) of rotaviruses. *J Gen Virol* (66 ( Pt 9)): 1889-900.
36. Hu, Y.-C. and Bentley, W.E., **2000**. A kinetic and statistical-thermodynamic model for baculovirus infection and virus-like particle assembly in suspended insect cells. *Chem Eng Sci* (55): 3991-4008.
37. Zlotnick, A., **1994**. To build a virus capsid. An equilibrium model of the self assembly of polyhedral protein complexes. *J Mol Biol* 1 (241): 59-67.
38. Grimes, J.M., *et al.*, **1998**. The atomic structure of the bluetongue virus core. *Nature* 6701 (395): 470-8.
39. Grimes, J.M., *et al.*, **1997**. An atomic model of the outer layer of the bluetongue virus core derived from X-ray crystallography and electron cryomicroscopy. *Structure* 7 (5): 885-93.
40. Mathieu, M., *et al.*, **2001**. Atomic structure of the major capsid protein of rotavirus: implications for the architecture of the virion. *Embo J* 7 (20): 1485-97.
41. Zlotnick, A., **2003**. Are weak protein-protein interactions the general rule in capsid assembly? *Virology* 2 (315): 269-74.



# Chapter 5

## Model-based assessment of multiple infection strategies

This chapter was based on the following manuscripts:

António Roldão, Helena L. A. Vieira, Paula M. Alves, R. Oliveira, Manuel J.T. Carrondo (2006). "Intracellular dynamics in rotavirus-like particles production: evaluation of multigene and monocistronic infection strategies", *Process Biochemistry*, 41(10): 2188-2199.

and

António Roldão, Helena L.A. Vieira, Annie Charpilienne, Didier Poncet, Polly Roy, Manuel J.T. Carrondo, Paula M. Alves and R. Oliveira (2007). "Modeling Rotavirus-Like Particles production in a Baculovirus Expression Vector System: infection kinetics, baculovirus DNA replication, mRNA synthesis and protein production", *Journal of Biotechnology*, 128(4): 875-894.

## Summary

Rotavirus disease is the prime cause of severe gastrointestinal illness in children, with an incidence estimated at 111 million episodes and a total of 440.000 deaths per year in children with less than 5 years of age. Rotavirus-like particles (RLPs) are excellent vaccine candidates against rotavirus infection since they are non-infectious, highly immunogenic, amenable to large-scale production and safer to produce than attenuated or inactivated vaccines. This work focuses on the analysis and modelling of the major events taking place inside *Spodoptera frugiperda* Sf-9 cells infected by recombinant baculovirus that may be critical for the expression of rotavirus proteins (vp), vp2, vp6 and vp7. For model calibration and validation, several experiments were performed adopting either a co-infection or a single-infection strategy using monocistronic, bicistronic or tricistronic baculovirus vectors. A characteristic viral DNA (vDNA) replication rate of  $0.19 \pm 0.01 \text{h}^{-1}$  was obtained irrespective of the infection strategy and the synthesis of progeny virus was found to be negligible in comparison to intracellular vDNA concentrations. The timeframe for vDNA, mRNA and protein synthesis tends to decrease with increasing multiplicity of infection due to the metabolic burden effect. The rate of protein synthesis varies due to codon usage and/or size of protein and also the time for maturation. The model exhibits acceptable prediction power of the dynamics of intracellular vDNA replication, mRNA synthesis and protein production for the three proteins involved. This model is intended to be the basis for future RLP process optimization and also a means to evaluating different baculovirus constructs for RLP production.

## CONTENTS

<b>1. Introduction .....</b>	<b>152</b>
<b>2. Model formulation .....</b>	<b>154</b>
2.1. Process description .....	154
2.2. Modelling assumptions .....	155
2.3. Baculovirus adsorption .....	156
2.4. Infected and healthy cell population .....	158
2.5. Replication of vDNA.....	159
2.6. Synthesis of mRNA.....	161
2.7. Viral protein synthesis.....	161
2.8. Temporal control.....	162
2.9. RLPs assembly.....	164
<b>3. Materials and methods .....</b>	<b>164</b>
3.1. Cell culture and media .....	164
3.2. Baculoviruses design and infection strategies.....	165
3.3. Infection procedure .....	165
3.4. Nucleic acid quantification .....	166
3.5. Viral protein analysis.....	167
3.6. RLPs purification, quantification and characterization.....	167
3.7. Parameter estimation algorithm.....	168
<b>4. Results and discussion .....</b>	<b>169</b>
4.1. Model calibration and validation .....	169
4.2. Effect of viral infection on cell death rate.....	170
4.3. The dynamics of intracellular vDNA .....	171
4.4. Virus budding.....	173
4.5. mRNA synthesis .....	174
4.6. Protein synthesis .....	176
4.7. Analysis of protein dynamics .....	178
<b>5. Conclusions.....</b>	<b>179</b>
<b>6. Acknowledgements .....</b>	<b>181</b>
<b>7. References.....</b>	<b>181</b>

1. INTRODUCTION

According to UNICEF/WHO report of October 2009 focused on diarrheal disease, the second leading killer of children under 5, rotavirus disease is responsible for more than 600,000 deaths in infants and children worldwide annually (source: <http://www.unicef.org>). More than 82% of deaths occur in developing countries where medical care centers are scarce and in most cases inexistent. In developed countries, the most prominent burden is economic (Table 14).

**Table 14.** The burden caused by rotavirus disease.

	Portugal	Europe	USA
Number of cases	66.565	2.800.042	3.000.000
Doctor visits	16.641	700.000	400.000
Hospitalizations	2.080	87.000	100.000
Cost Hospitalization (millions of euros)	3	125,48	200,4
Cost Hospitalization / child (euros)	1.442	1.442	2.004
Total cost (millions of euros)	11,4	476	760
% of hospitalization due to gastroenteritis	40%	40-50%	40-50%
Number of deaths	3	231	20-100

Sources: Sociedade Portuguesa Pediatria, US Centers for Disease Control and Prevention, US National Institute of Allergy and Infectious Diseases

Rotavirus disease is caused by rotavirus, a non-enveloped double-stranded RNA virus with icosahedral geometry and composed by six structural proteins (vp1, vp2, vp3, vp4, vp6 and vp7) and five non-structural proteins (NSP1 to NSP5).

The first vaccine against rotavirus disease, RotaShield<sup>®</sup>, was released in the US in August 1998 by Merck & Co, Inc. This simian tetravalent vaccine was withdrawn from the market by its manufacturer one year later after the discovery of a rare association with intussusception in vaccinated children [1]. In 2006, two vaccines based on live attenuated viruses, Rotateq<sup>®</sup> from Merck & Co, Inc. and Rotarix<sup>®</sup> from GlaxoSmithKline Inc. [2; 3], were licensed in the US and Europe yielding, to date, positive results [4; 5]. Nonetheless, the possibility of virus mutation and reversion to virulence,



reassortment and serious adverse events continues associated with this type of vaccines.

To overcome such challenges, the use of virus-like particles (VLPs) as candidate vaccines may constitute an attractive alternative. These multimeric protein structures can mimic the structure of authentic viruses, preserving the native antigenic conformation of the immunogenic proteins, and do not possess the virus DNA, potentially yielding a safer and cheaper vaccine than those based on attenuated viruses. Moreover, VLPs can be engineered to contain structural proteins corresponding to different virus serotypes, thus providing a broader protection. With new VLP vaccines reaching the market, such as the human papillomavirus vaccine, this technology must be considered valid for the development of a rotavirus-like particle (RLP) based vaccine against rotavirus disease.

RLPs are spherical shaped particles composed by three viral proteins (vp) of rotavirus arranged in a triple layered structure: the innermost layer is composed by 60 dimers of vp2 (102.7 kDa) [6]; the middle shell is formed by 260 trimers of vp6 (44.9 kDa) [7] and the third, outer layer is composed by 260 trimers of glycoprotein vp7 (37.2 kDa) [7]. RLPs are normally expressed in the baculovirus/insect cell system [8].

Due to the complexity of RLP structure, the achievement of economical volumetric productivities is extremely difficult; previous studies reported wasteful accumulation of unassembled proteins and formation of incomplete RLPs [9]. In fact, less than 12% of expressed proteins find their way into correctly assembled RLPs probably due to incorrect stoichiometric ratios or inadequate thermodynamic aggregation conditions.

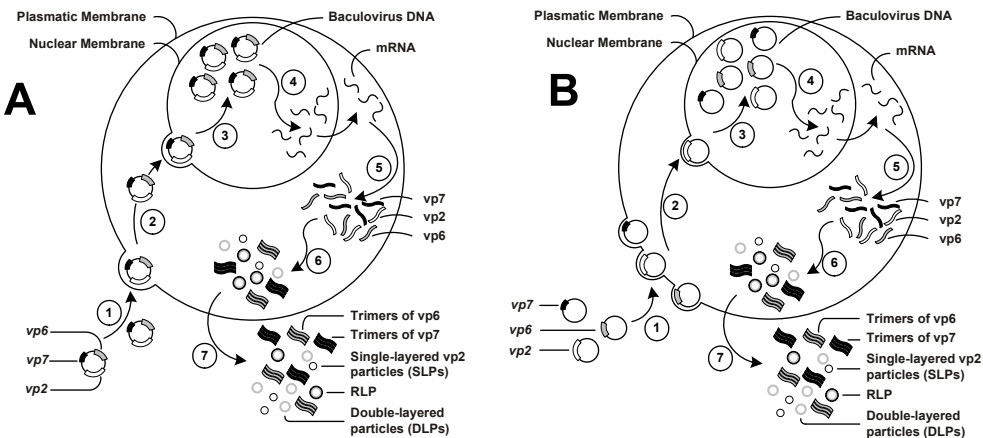
In this work we propose a deterministic mathematical model of the baculovirus/insect cell system for the production of viral proteins and RLPs.

The model addresses the most critical steps of baculovirus internalization and trafficking, and gene expression mechanisms. The main objective of this modelling study is to understand and quantify the effect on viral protein synthesis and RLP formation of different infection strategies, and evaluate the effect of different promoters.

## 2. MODEL FORMULATION

### 2.1. Process description

The main steps involved in RLP production are represented schematically in Figure 26 for single (S2/6/7) and co-infection (Co2+6+7) strategies. In co-infection, three monocistronic recombinant baculoviruses, each one enclosing the genes coding for vp2, vp6 or vp7, were used. In single-infection, tricistronic recombinant baculoviruses enclosing the genes coding for the three viral proteins simultaneously were used.



**Figure 26.** Schematic representation of the RLPs production process for single-infection S2/6/7 (A) and co-infection Co2+6+7 (B) strategies: 1. virus adsorptive endocytosis; 2. virus trafficking; 3. vDNA replication; 4. vDNA transcription into the corresponding mRNA; 5. protein synthesis; 6. RLPs assembly and 7. RLPs release to the extracellular medium.

Independently of the strategy, the triggering event is the baculovirus binding to the plasmatic membrane and their entry to the cell by adsorptive endocytosis [10]. Step 2 consists in virus trafficking into the cell nucleus. Baculoviruses are first enclosed within endosomes in the cytoplasmic space. As the pH inside the endosomes drops, the viruses' nucleocapsids, containing the viral genes, are released to the cytoplasm and migrate to the cell nucleus. Then, the nucleocapsids bind to the nucleus membrane, thereby delivering the genes into the cell nucleus. Once inside the nucleus, viral genes are expressed in a well orchestrated cascade that is normally divided in three stages: early, late (6-15 hpi) and very late (>15 hpi) [11]. Viral DNA (vDNA) replication (step 3) commences at the beginning of the late phase. The vDNA is transcribed into the corresponding mRNA during the very late phase, under the control of the *polh* promoter or under the control of the *p10* promoter, depending on the construct of the recombinant baculovirus used (step 4). Synthesized mRNA molecules leave the cell nucleus and migrate to ribosomes where viral proteins are synthesized (step 5). These newly synthesized proteins (vp2, vp6 and vp7) may assemble into trimers of vp6 and vp7, single-layered vp2 particles, double-layered particles of vp2 and vp6 (DLPs) and correctly formed triple-layered RLP (step 6). This complex mixture of structures is released to the extracellular medium in the final step (step 7).

## 2.2. Modelling assumptions

A set of simplifying hypothesis were assumed:

- Excess of nutrients: the extracellular nutrients do not limit the synthesis of virus related products;
- Synchronous infection: in a batch culture of insect cells with a multiplicity of infection (MOI – number of virus *per* cell) higher than 2 virus.cell<sup>-1</sup>, the ensuing infection process will be essentially

synchronous, i.e., all cells will go through the infection cycle simultaneously [12]. Such MOIs ensure that all cells are infected, preventing cell division from competing for nutrient resources needed for virus and protein synthesis [13];

- Infection kinetics is independent of the viral gene in the recombinant baculovirus. Power and Nielsen (1996) showed that recombinant baculovirus carrying different foreign genes present similar infection kinetics in *Sf-9* suspension cultures;
- Virus trafficking, encompassing attachment, endocytosis, uncoating and transport to the nucleus are independent of intracellular state;
- One-half of internalized viruses are routed to lysosomes for degradation [13];
- The genes coding for each viral protein do not affect vDNA replication. The size of the monocistronic, bicistronic and tricistronic baculovirus constructs is similar; thus, vDNA replication kinetics is assumed to be the same for the different constructs;
- Negligible virus budding and release: our data suggests that for MOIs  $> 2 \text{ virus.cell}^{-1}$  the rate of virus budding is much lower than the vDNA replication rate. This hypothesis has also been defended by Rosinski *et al.* (2002);
- Negligible vp2-vp6-vp7 interaction: the interaction between the different proteins during synthesis is unknown;
- Stoichiometric RLP assembly: the final RLPs titer is determined by the limiting protein.

### 2.3. Baculovirus adsorption

According to Dee *et al.* (1995, 1997), the depletion of extracellular virus due to binding to cell surface is:

$$\frac{dV_j}{dt} = -k_a (N_i + N_u) V_j \quad \text{Eq. 5.1}$$

with  $V_j$  the concentration of extracellular recombinant baculovirus coding for  $vp_j$  (DNA.ml<sup>-1</sup>), where subscript index  $j=2,6,7$  representing  $vp_2$ ,  $vp_6$  and  $vp_7$  respectively,  $k_a$  the attachment rate (ml.cell<sup>-1</sup>.h<sup>-1</sup>), considered here to be independent of the gene coding for the viral protein,  $N_i$  is the concentration of infected cells (cell.ml<sup>-1</sup>),  $N_u$  is the concentration of uninfected cells, (cell.ml<sup>-1</sup>) and  $t$  is time post-infection (hpi). The attachment rate is defined as follows [14]:

$$k_a = k_f(\alpha R) \quad \text{Eq. 5.2}$$

with  $k_f$  the intrinsic forward rate constant for the binding of a single viral attachment protein to a cell membrane receptor,  $\alpha$  is the number of attachment proteins per virus, ~1000 for baculovirus [15; 16] and  $R$  the number of free surface receptors per cell. In typical infection conditions (MOI < 50 virus.cell<sup>-1</sup>), the number of surface receptors per cell is normally in great excess,  $R \sim 11000$  for *Sf*-21 cells [13], thus the attachment rate  $k_a$  may be assumed to be time-invariant. According to Dee and Shuler (1997), the attachment rate constant variation between baculovirus infection for *Sf*-21 and *Sf*-9 cells in suspended cultures is meaningless; thus, a  $k_a$  of  $7.8 \times 10^{-8}$  ml.cell<sup>-1</sup>.h<sup>-1</sup> was adopted for this study [13]. The adsorption of baculoviruses to *Sf*-9 cells was also studied by Power *et al.* (1994) for MOIs < 10 virus.cell<sup>-1</sup> in suspended cultures. It was verified that the attachment rate is independent of MOI, but slightly dependent of the cell density, as it is reported in the present study.

For co-infection strategies (experiments 5, 6 and 8 in Table 15), the total extracellular virus concentration at a given instant,  $V_t$  (DNA.ml<sup>-1</sup>), is given by the sum of individual extracellular virus concentration:

$$V_t = \sum_j V_j \quad \text{Eq. 5.3}$$

For single-infection strategies (experiments 1 to 4, 7 and 9 in Table 15), the total extracellular virus concentration is equal to the concentration of any extracellular virus carrying viral gene  $j$ :

$$V_t = V_j, \quad \forall_j \quad \text{Eq. 5.4}$$

**Table 15.** Experimental setup for protein expression.

<b>Cell Culture Conditions</b>	250 ml Spinner @ CCI ~ $1 \times 10^6$ cell.ml <sup>-1</sup>
<b>Measured times</b>	0/5/10/15/20/25/30/35/40/45/57/72/84/96/120/144 hpi
<b>Calibration Experiments</b>	Exp. 1: S2 - BacRF2A - MOI = 5 Exp. 2: S6 - BacVP6C - MOI = 5 Exp. 3: S7 - BacRF7 - MOI = 5 Exp. 4: S2/6/7 - BRV RF VLP2/6/7 - MOI = 5 Exp. 5: Co2+6+7 - BacRF2A + BacVP6C + BacRF7 - MOI = 15
<b>Validation Experiments</b>	Exp. 6: Co2+6 - BacRF2A + BacVP6C - MOI = 10 <sup>A</sup> Exp. 7: S2/6 - VP2(RF)/VP6(RF) VLPs - MOI = 5 <sup>A</sup> Exp. 8: Co2/6+7 - VP2(RF)/VP6(RF) VLPs + BacRF7 - MOI = 10 <sup>A</sup> Exp. 9: S2/6/7 - BRV RF VLP2/6/7 - MOI = 15 <sup>A</sup>

<sup>A</sup> - mRNA data not available

## 2.4. Infected and healthy cell population

The material balance equation of infected cells has two terms. The first term accounts for the increase in infected cells concentration due to binding of baculoviruses to uninfected cells,  $N_u$  (cell.ml<sup>-1</sup>). The second term represents the cells death rate:

$$\frac{dN_i}{dt} = k_a N_u V_t \left( \frac{1}{MOI} \right) - k_d N_i \quad \text{Eq. 5.5}$$

The viability of infected cells exhibits two phases as previously described in the literature [17]. The first phase is characterized by a slight decrease in infected cells concentration related to the intrinsic cell death rate. In the second phase, cell lysis is considerably faster due to viral infection. Accordingly, the cell death rate,  $k_d$  (h<sup>-1</sup>) has two terms:

$$k_d = \begin{cases} k_{d1} & t < \delta_D \\ k_{d1} + k_{d2} & t \geq \delta_D \end{cases} \quad \text{Eq. 5.6}$$

with  $k_{d1}$  the intrinsic cell death rate, which is equal to  $0.0008 \text{ h}^{-1}$  [17; 18],  $k_{d2}$  the increase in cell death rate due to infection ( $\text{h}^{-1}$ ) and  $\delta_D$  (hpi) the time instant when the cell death rate increases. Both  $k_{d2}$  and  $\delta_D$  (hpi) show a relationship with the extent of host cells infection; namely, the parameter  $k_{d2}$  increases with the number of viruses infecting the cells:

$$k_{d2} = \begin{cases} k^* & DNA_{total} \leq 10 \\ k^* \log(DNA_{total}) & DNA_{total} > 10 \end{cases} \quad \text{Eq. 5.7}$$

with  $k^*$  ( $\text{h}^{-1}$ ) the increase in cell death rate correspondent to 10 intracellular vDNA copies and  $DNA_{total}$  the total number of intracellular vDNA copies (see equations 5.11 and 5.12). The calculation of  $\delta_D$  is discussed at the end of this section.

The material balance equation of uninfected cells,  $N_u$  ( $\text{cell.ml}^{-1}$ ), comprises two terms. The first term represents the “conversion” of uninfected cells into infected cells due to virus binding. The second term reports the intrinsic cell death rate:

$$\frac{dN_u}{dt} = -k_{aVt} N_u V_t \left( \frac{1}{MOI} \right) - k_{d1} N_u \quad \text{Eq. 5.8}$$

## 2.5. Replication of vDNA

The dynamics of vDNA in the cell nucleus,  $DNA_j^{nuc}$  ( $\text{DNA.cell}^{-1}$ ) is given by the following material balance equation:

$$\begin{aligned} \frac{dDNA_j^{nuc}}{dt} = & \eta_{traf} k_{aVj} (t - \tau_{traf}) \left( 1 + \frac{N_u}{N_i} \right) + k_{RDNA} DNA_j^{nuc} \\ & \times f_{DNA,rep}(t, \delta_{DNA,low}, \delta_{DNA,high}) \end{aligned} \quad \text{Eq. 5.9}$$

The first term accounts for the transport of vDNA from extracellular virus into the cell nucleus. Parameter  $\eta_{\text{traf}}$  is the trafficking efficiency, which is the fraction of internalized genes that manage to enter in the cell nucleus. Considering that about half of internalized baculovirus is degraded, most likely in lysosomes which contain the hydrolytic enzymes able to degrade the virus rapidly [13; 19], then  $\eta_{\text{traf}}$  is equal to 0.5. Since virus trafficking within the cell takes some time, this transport term has a delay  $\tau_{\text{traf}}$  (h). According to Dee and Shuler (1997), the mean time for trafficking is as follows: 43 min for endocytosis, 40 to 60 min for uncoating and 6 to 25min for transport to nucleus, which sum  $\tau_{\text{traf}} = 1.8$  h.

The second term is the vDNA replication kinetics, with  $k_{\text{RDNA}}$  ( $\text{h}^{-1}$ ) the first order replication constant. The time-dependent function  $f_{\text{DNA,rep}}(t, \delta_{\text{DNA,low}}, \delta_{\text{DNA,high}})$  defines a time interval  $[\delta_{\text{DNA,low}}, \delta_{\text{DNA,high}}]$  for vDNA replication, accounting also for the effect of the metabolic decay on vDNA replication (see equation 5.17).

The total number of vDNA copies inside the cell,  $\text{DNA}_j^T$  ( $\text{DNA} \cdot \text{cell}^{-1}$ ), is the sum of vDNA within the nucleus and vDNA in the cytoplasm:

$$\frac{d\text{DNA}_j^T}{dt} = \eta_{\text{traf}} k_a V_j \left( 1 + \frac{N_u}{N_i} \right) + k_{\text{RDNA}} \text{DNA}_j^{\text{nuc}} \times f_{\text{DNA,rep}}(t, \delta_{\text{DNA,low}}, \delta_{\text{DNA,high}}) \quad \text{Eq. 5.10}$$

The first term accounts for the internalization of extracellular virus while the second corresponds to the degradation kinetics in lysosomes. The third term reports the vDNA replication kinetics. The total number of vDNA copies inside the cell in the co-infection experiments (experiments 5, 6 and 8 in Table 15) is given by the sum of individual intracellular vDNA copies:

$$\text{DNA}_{\text{total}} = \sum_j \text{DNA}_j^T \quad \text{Eq. 5.11}$$



For single-infection experiments (experiments 1 to 4, 7 and 9 in Table 15), the total number of vDNA copies inside the cell corresponds to the individual intracellular vDNA copies:

$$DNA_{total} = DNA_j^T \quad \forall j \quad \text{Eq. 5.12}$$

## 2.6. Synthesis of mRNA

The transcription of the genes coding for **vp<sub>2</sub>**, **vp<sub>6</sub>** or **vp<sub>7</sub>** into the corresponding mRNA, under the control of a very late promoter (either *polh* or *p10* promoter), follows first order kinetics on the corresponding vDNA templates:

$$\frac{dRNA_j}{dt} = k_{SRNA,j} DNA_j^{nuc} f_{VP}(t, \delta_{VP,low}, \delta_{VP,high}) - k_{DRNA,j} RNA_j \quad \text{Eq. 5.13}$$

with **RNA<sub>j</sub>** (RNA.cell<sup>-1</sup>) the intracellular concentration of mRNA, **k<sub>SRNA,j</sub>** (h<sup>-1</sup>) the first order transcription rate and **k<sub>DRNA,j</sub>** the first order mRNA degradation rate (h<sup>-1</sup>). The transcription of vDNA and the synthesis of viral protein obey to temporal control mechanisms. Therefore, a time interval for mRNA synthesis [**δ<sub>VP,low</sub>**, **δ<sub>VP,high</sub>**] is defined by the function **f<sub>VP</sub>** (**t**, **δ<sub>VP,low</sub>**, **δ<sub>VP,high</sub>**) (see equation 5.18).

## 2.7. Viral protein synthesis

The kinetics of structural viral protein synthesis, **VP<sub>j</sub>** (μg.ml<sup>-1</sup>), was defined with Michaelis-Menten kinetics on mRNA:

$$\frac{dVP_j}{dt} = k_{VP,j} \frac{RNA_j}{K_{RNA} + RNA_j} f_{VP}(t, \delta_{VP,low}, \delta_{VP,high}) N_i \quad \text{Eq. 5.14}$$

where **k<sub>VP,j</sub>** (μg.cell<sup>-1</sup>.h<sup>-1</sup>) is the maximum **VP<sub>j</sub>** synthesis rate and **K<sub>RNA</sub>** (RNA.cell<sup>-1</sup>) the half-saturation constant for intracellular mRNA. The

intracellular protein content,  $\mathbf{VP}_j^{\text{int}}$  ( $\mu\text{g}\cdot\text{cell}^{-1}$ ) is given by:

$$\frac{dVP_j^{\text{int}}}{dt} = k_{VP,j} \frac{RNA_j}{K_{RNA} + RNA_j} f_{VP}(t, \delta_{VP,low}, \delta_{VP,high}) \quad \text{Eq. 5.15}$$

Note that both  $\mathbf{VP}_j$  and  $\mathbf{VP}_j^{\text{int}}$  account for assembled and unassembled proteins altogether.

The maximum viral protein synthesis rate,  $\mathbf{k}_{VP,j}$ , shows an indirect relationship with the intracellular vDNA concentration probably via the effect of translation enhancing factors coded by early and/or late genes. Here, such effect has been expressed as follows:

$$k_{VP,j} = \begin{cases} k_{VP,j}^* & DNA_{total} \leq 10 \\ k_{VP,j}^* \log(DNA_{total}) & DNA_{total} > 10 \end{cases} \quad \text{Eq. 5.16}$$

with  $\mathbf{k}_{VP,j}^*$  ( $\mu\text{g}\cdot\text{cell}^{-1}\cdot\text{h}^{-1}$ ) the maximum  $\mathbf{VP}_j$  synthesis rate correspondent to 10 intracellular vDNA copies and  $\mathbf{DNA}_{total}$  the total number of intracellular vDNA copies (see equations 5.11 and 5.12).

## 2.8. Temporal control

The expression of viral genes occurs in an ordered cascade of events, whereby the early genes transactivate directly or indirectly the late genes, which in turn transactivate the very late genes. This sequence of events results in well defined time intervals for vDNA replication and transcription, and for protein expression that may be expressed as functions of the intracellular viral load [18; 20; 21].

The vDNA replication, which commences in the late phase at  $\delta_{DNA,low} = 6$  hpi [11], is strongly affected during the overall process since infected cells gradually lose their ability to synthesize vDNA. In their segregated

population framework, Licari and Bailey (1992) expressed this metabolic decay by a linear decay in time. The same concept was adopted here to define the time window for vDNA replication:

$$f_{DNA,rep}(t, \delta_{DNA,low}, \delta_{DNA,high}) = \begin{cases} 0 & t < \delta_{DNA,low} \\ 1 - \frac{t - \delta_{DNA,low}}{\delta_{DNA,high} - \delta_{DNA,low}} & \delta_{DNA,low} \leq t < \delta_{DNA,high} \\ 0 & t \geq \delta_{DNA,high} \end{cases} \quad \text{Eq. 5.17}$$

The expression of viral proteins which are under the control of very late promoters, *polh* or *p10*, starts approximately at  $\delta_{VP,low} = 15$  hpi [11; 18; 20; 22]. Hu and Bentley (2000, 2001) and Licari and Bailey (1992) considered a linear time decay function for viral protein production. Here, a similar approach was adopted:

$$f_{VP}(t, \delta_{VP,low}, \delta_{VP,high}) = \begin{cases} 0 & t < \delta_{VP,low} \\ 1 - \frac{t - \delta_{VP,low}}{\delta_{VP,high} - \delta_{VP,low}} & \delta_{VP,low} \leq t < \delta_{VP,high} \\ 0 & t \geq \delta_{VP,high} \end{cases} \quad \text{Eq. 5.18}$$

Interestingly, our data suggests the existence of three critical events practically coincident in time: 1. the increase in cell death rate; 2. the halt in vDNA replication and 3. the halt in viral protein expression. These major events are somehow interconnected in what appears to be an influence of intracellular viral load. This specific moment in time has been previously expressed as a function of the number of viruses infecting the host cell [18; 20; 21]. Here, this time instant is expressed as a dynamic function of the intracellular vDNA:

$$\delta_{DNA,high} = \delta_{VP,high} = \delta_D = \begin{cases} \delta^* & DNA_{total} \leq 10 \\ \frac{\delta^*}{\log(DNA_{total})} & DNA_{total} > 10 \end{cases} \quad \text{Eq. 5.19}$$

with  $\delta^*$  (hpi) the critical time instant correspondent to 10 intracellular vDNA copies.

## 2.9. RLPs assembly

The assembly of RLPs is complex to model since they involve three viral proteins. During the assembly process, many different stable subunits are possible and there is yet limited knowledge of the details of RLPs assembly in insect cells [23]. A simpler approach was adopted, based on protein stoichiometric ratios between  $\mathbf{vp}_6$  and  $\mathbf{vp}_2$  ( $Y_{6/2}$ ), and between  $\mathbf{vp}_7$  and  $\mathbf{vp}_2$  ( $Y_{7/2}$ ), and on their comparison with the composition of correctly assembled particles, as a mean to evaluate the dynamics of RLPs assembly for possible process limitations analysis. The relative mass composition of  $\mathbf{vp}_6/\mathbf{vp}_2$  and  $\mathbf{vp}_7/\mathbf{vp}_2$  in correctly assembled particles is  $Y_{6/2} = 2.8$  (w/w) and  $Y_{7/2} = 2.4$  (w/w) respectively.

## 3. MATERIALS AND METHODS

### 3.1. Cell culture and media

*Spodoptera frugiperda* Sf-9 and Sf-21 cell lines, obtained from American Type Culture Collection (ATCC, Washington D.C., US), were cultivated in serum free media SF900II (Gibco, Glasgow, UK) at 27 °C in 500 ml erlenmeyers (50 ml working volume) (DURAN, Mainz, Germany) at 90 rpm and routinely diluted every four days using an inoculum of  $0.3 \times 10^6$  cell.ml<sup>-1</sup>. Cell concentration and viability were determined by counting cells on a Fuchs-Rosenthal haemocytometer (Brandt, Wertheim/Main, Germany) using the trypan blue exclusion dye (Merck, Darmstadt, Germany) method or using a Casy<sup>®</sup>1 Cell counter plus Analyzer System Model TTC (Innovatis AG, Reutlingen, Germany).

### 3.2. Baculoviruses design and infection strategies

Different baculovirus constructs were used: 1. monocistronic recombinant baculoviruses BacRF2A [6], BacVP6C [24] and BacRF7 enclosing the genes coding for  $vp_2$ ,  $vp_6$  and  $vp_7$  respectively; 2. bicistronic recombinant baculoviruses VP2(RF)/VP6(RF) VLPs with genes coding for  $vp_2$  and  $vp_6$  simultaneously; 3. tricistronic recombinant baculoviruses BRV RF VLP2/6/7 whose genes encode the three viral proteins simultaneously. All genes were under the control of the *polh* promoter with the exception of the gene coding for  $vp_7$  in the tricistronic recombinant baculovirus that is under the control of *p10* promoter.

Two infection strategies were evaluated: 1. *Co-infection* – cells were infected with the two monocistronic baculovirus whose genes encode  $vp_2$  and  $vp_6$  (Co2+6) (Exp. 6 in Table 15), with the three monocistronic baculoviruses (Co2+6+7) (Exp. 5 in Table 15) or with the bicistronic baculovirus plus the monocistronic baculovirus enclosing the gene coding for  $vp_7$  (Co2/6+7) (Exp. 8 in Table 15); 2. *Single-infection* – cells were infected with monocistronic baculoviruses enclosing the genes coding for  $vp_2$  (S2),  $vp_6$  (S6) or  $vp_7$  (S7) individually (Exp. 1, 2, 3 in Table 15), with the bicistronic baculovirus (S2/6) (Exp. 7 in Table 15) or with the tricistronic baculovirus (S2/6/7) (Exp. 4, 9 in Table 15).

### 3.3. Infection procedure

Recombinant baculoviruses were amplified by infecting Sf-9 or Sf-21 cells at a cell concentration of  $1 \times 10^6$  cell.ml<sup>-1</sup> and MOI of 0.1 virus.cell<sup>-1</sup> in 250 ml working volume spinner flasks (Wheaton Science, New Jersey, US) stirred at 150 rpm and at 27 °C. Baculovirus were harvested at 60-70% of cell viability, around 96-120 hpi, by centrifugation at  $200 \times g$  for 15 min at 4 °C. Titers were assessed using several titration methods (see Roldão *et al.*

2009). For viral protein and RLP production, *Sf*-9 cells were infected at a cell concentration of  $1 \times 10^6$  cell.ml<sup>-1</sup> in 250 ml working volume spinner flasks (Wheaton Science, New Jersey, US) stirred at 150 rpm and at 27 °C.

### **3.4. Nucleic acid quantification**

High pure viral nucleic acid kit and high pure RNA isolation kit (Roche Diagnostics, Mannheim, Germany) were used for extraction and purification of recombinant baculovirus DNA and RNA, respectively. Purified samples were stored at -20 °C for DNA and -80 °C for RNA. RNA integrity was evaluated by agarose gel electrophoresis and total RNA content assessed by spectral absorption of UV light at 260 and 280 nm using a GeneQuant II unit (Pharmacia Biotech, Cambridge, UK). Synthesis of cDNA from purified RNA samples was performed using first strand cDNA synthesis kit (Roche Diagnostics, Mannheim, Germany) and sequence-specific primers designed for the genes coding for vp<sub>2</sub>, vp<sub>6</sub> and vp<sub>7</sub>. The final cDNA was stored at -20 °C. Purified DNA and cDNA were analyzed by real-time quantitative PCR in LightCycler<sup>®</sup>1.5 instrument (Roche Applied Science, Indianapolis, US) using Fast start DNA master SYBR Green I kit (Roche Diagnostics, Mannheim, Germany) and sequence-specific primers for the genes coding for vp<sub>2</sub>, vp<sub>6</sub> and vp<sub>7</sub> (see Vieira *et al.* (2005)). Based on the standard calibration curve of cross points vs log concentrations of the purified DNA standard with known concentration (virus.ml<sup>-1</sup>), the content in DNA of each sample was assessed. Degradation of mRNA was evaluated by adding Actinomycin D (1 µg.ml<sup>-1</sup>) (Sigma-Aldrich, St. Louis, US) to cells at 42 hpi, maximum pick of mRNA, to inhibit vDNA transcription. Then, cells were harvested, followed by RNA purification, cDNA synthesis and PCR quantification.

### 3.5. Viral protein analysis

Protein analysis was performed by Western blot. Briefly, cell culture samples were diluted in NuPAGE<sup>®</sup> LDS sample buffer (Invitrogen, California, US) and incubated for 10 min at 70 °C. Denatured proteins were separated by gel electrophoresis using 1 mm NuPAGE<sup>®</sup> 4-12% Bis-Tris pre-cast polyacrylamide gels in NuPAGE<sup>®</sup> MOPS SDS running buffer (Invitrogen, California, US). Proteins were transferred to a Hybond<sup>™</sup>-C extra nitrocellulose membrane (Amersham Biosciences, Fairfield, US) using a semi-dry transfer unit (GE Healthcare, Uppsala, Sweden) and probed using antibodies specific to target the proteins of interest (*vp2*, *vp6* and *vp7*). Different primary antibodies were used: 1) mouse monoclonal anti-*vp2* IgG; 2) mouse monoclonal anti-*vp6* IgG; 3) mouse monoclonal [511] anti-rotavirus (Abcam, Cambridge, UK); 4) goat polyclonal anti-rotavirus (Abcam, Cambridge, UK); 5) rabbit polyclonal anti-rotavirus IgG. Blots were developed after incubation with alkaline phosphatase conjugated anti-mouse IgG, anti-goat IgG or anti-rabbit IgG (all from Sigma-Aldrich, St. Louis, US) using 1-step<sup>™</sup> NBT/BCIP blotting detection reagents (Pierce, Rockford, US). Protein concentration was estimated by densitometry analysis of scanned images using the open source Image J software version 1.41a (<http://rsb.info.nih.gov/ij/>) or ImageQuant<sup>®</sup> software for Microsoft<sup>®</sup> Windows NT<sup>®</sup> (Molecular Dynamics, Inc., US, 1998).

### 3.6. RLPs purification, quantification and characterization

The purification of RLPs was performed using *in-house* developed methods consisting in depth filtration, ultrafiltration and size exclusion chromatography as stepwise unit operations [25; 26].

The total protein content in purified RLPs was determined using the BCA protein quantification assay kit (96-well plate protocol) (Pierce, Rockford,

US). Characterization of RLPs (integrity and morphology) was performed by TEM as described elsewhere [26].

### 3.7. Parameter estimation algorithm

A kinetic parameter estimation program was developed in MATLAB<sup>TM</sup> (The MathWorks, Inc., US, 1994-2006). The program minimizes the residuals in the sense of least squares employing the Levenberg-Marquardt algorithm. The mathematical model equations 5.1 to 5.19 were integrated using the 4<sup>th</sup>/5<sup>th</sup> order Runge-Kutta solver (MATLABs' *ode45* function) modified to handle ordinary differential equations with time delays. The final residuals and Jacobian matrix served to calculate an approximation to the Hessian matrix, thereby assuming that the final solution is a local optimum. The Hessian matrix enabled to calculate the parameters covariance matrix and parameters 95% confidence intervals. The MATLABs' *nlparci* function was used for this calculation.

The accuracy of baculovirus stock titers is determined by the consistency and reproducibility of the titration method. Commonly applied methods present errors between 7% and 36% [27] that inherently affect the MOI determination for a specific experiment. The experimental error in the determination of the MOI is critical because it propagates exponentially in the calculation of the intracellular vDNA templates, thus having a major impact on the overall virus dynamics. To overcome this problem, for the purpose of kinetic parameters estimation, equations 5.1 to 5.7 were “bypassed” and the data of the first 5 hpi was not considered for parameter calibration. This allowed estimating more accurately the vDNA replication, transcription and translation parameters.



## 4. RESULTS AND DISCUSSION

### 4.1. Model calibration and validation

The calibration of the mathematical model used 5 experiments (Exp 1-5 in Table 15) while model validation used 4 additional experiments (Exp 6-9 in Table 15). Estimated parameters related to baculovirus infection and trafficking are presented in table 16 while estimated parameters related to viral protein expression are presented in table 17.

**Table 16.** Baculovirus infection and trafficking parameters.

Parameter	Value	Reference
$k_a$	$7.8 \times 10^{-8} \text{ml.cell}^{-1}.\text{h}^{-1}$	[19]
$k_{d1}$	$0.0008 \text{h}^{-1}$	[17; 18; 28]
$\tau_{traf}$	1.8h	[13]
$\eta_{traf}$	0.5	[13]
$k^*$	$0.0017 \pm 0.0003 \text{h}^{-1}$	[28]
$\delta^*$	$233 \pm 7 \text{h}$	[28]

**Table 17.** Viral protein expression parameters.

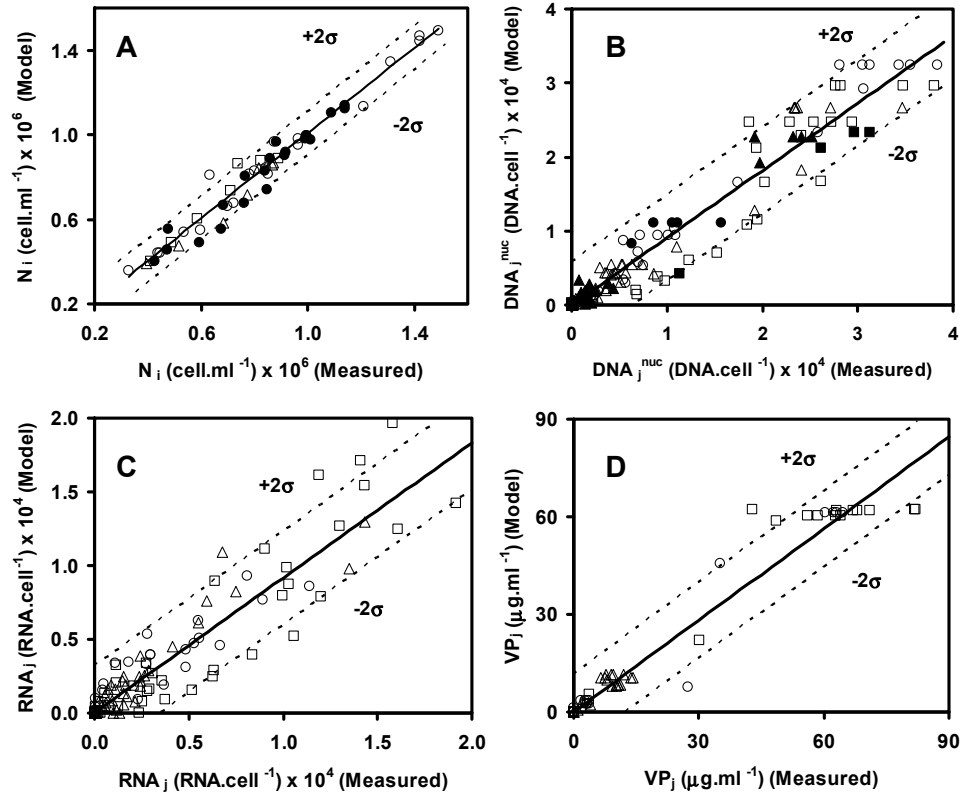
Parameter	vp2	vp6	vp7
$k_{RDNA} (\text{h}^{-1})$	$0.19 \pm 0.01$		
$\delta_{DNA,low} (\text{hpi})$	6		
$k_{SRNA_j} (\text{h}^{-1})$	<i>polh</i>	$0.12 \pm 0.01$	$0.16 \pm 0.02$
	<i>p10</i>	-	$0.46 \pm 0.01$
$k_{DRNA_j} (\text{h}^{-1})$	<i>polh</i>	$0.019 \pm 0.004$	$0.025 \pm 0.005$
	<i>p10</i>	-	$0.013 \pm 0.004$
$k^*_{VP_j} (\mu\text{g.cell}^{-1}.\text{h}^{-1}) \times 10^{-6}$	<i>polh</i>	$0.19 \pm 0.02$	$1.0 \pm 0.1$
	<i>p10</i>	-	$0.058 \pm 0.007$
$K_{RNA} (\text{RNA.cell}^{-1})$	$0.2 \times 10^4$ <sup>A</sup>		
$\delta_{VP,low} (\text{hpi})$	15		

<sup>A</sup> - fitted by trial and error

The parameters' estimates show high statistical confidence. Figure 27 shows model predictions against the corresponding measured variables for both the calibration and validation datasets.

The 95% confidence prediction errors are as follows:  $\mathbf{N}_i \pm 0.10 \times 10^6 \text{ cell.ml}^{-1}$ ,  $\mathbf{DNA}_j^{\text{nuc}} \pm 0.58 \times 10^4 \text{ DNA.cell}^{-1}$ ,  $\mathbf{RNA}_j \pm 0.32 \times 10^4 \text{ RNA.cell}^{-1}$  and  $\mathbf{VP}_j \pm$

$1.19 \times 10^1 \mu\text{g.ml}^{-1}$ . In the following sections, we analyze in detail the modelling results.



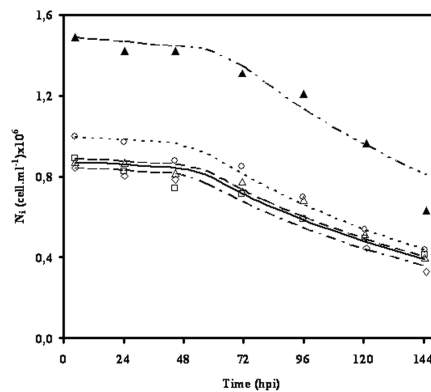
**Figure 27.** Model predictions over measured variables: infected cells concentration (A), intracellular vDNA (B) and mRNA (C) concentration, and total viral protein concentration (D). The full line (—) indicates the average model prediction and the dashed lines (---) represent 95% confidence prediction errors. Open ( $\triangle$ ,  $\square$  and  $\circ$ ) and closed ( $\blacktriangle$ ,  $\blacksquare$  and  $\bullet$ ) symbols denote data resulting from calibration and validation experiments respectively. Triangles, squares and circles represent data of genes coding for vp2, vp6 and vp7 respectively.

#### 4.2. Effect of viral infection on cell death rate

The intrinsic cell death rate of *Sf*-9 cells,  $k_{d1} = 0.0008\text{h}^{-1}$  is in agreement with previous studies [17; 18]. Upon viral infection, the increase in cell death rate, characterized here by  $k_{d2}$ , exhibits some variability depending

on the viral loading. For example, the  $k_{d2}$  is reportedly higher in the Co2+6+7 strategy (MOI = 15 virus.cell<sup>-1</sup>) than in the S2, S6, S7 and S2/6/7 experiments in consequence of the higher MOI used (MOI = 5 virus.cell<sup>-1</sup>), bringing about a more rapid decrease in cell viability (see for instance [29]).

Higher MOI leads to higher packing of baculoviruses inside the cell inducing faster cell death through apoptosis. This effect on the  $k_{d2}$  is well described by the empirical function defined in equation 5.7, as demonstrated in figure 28.

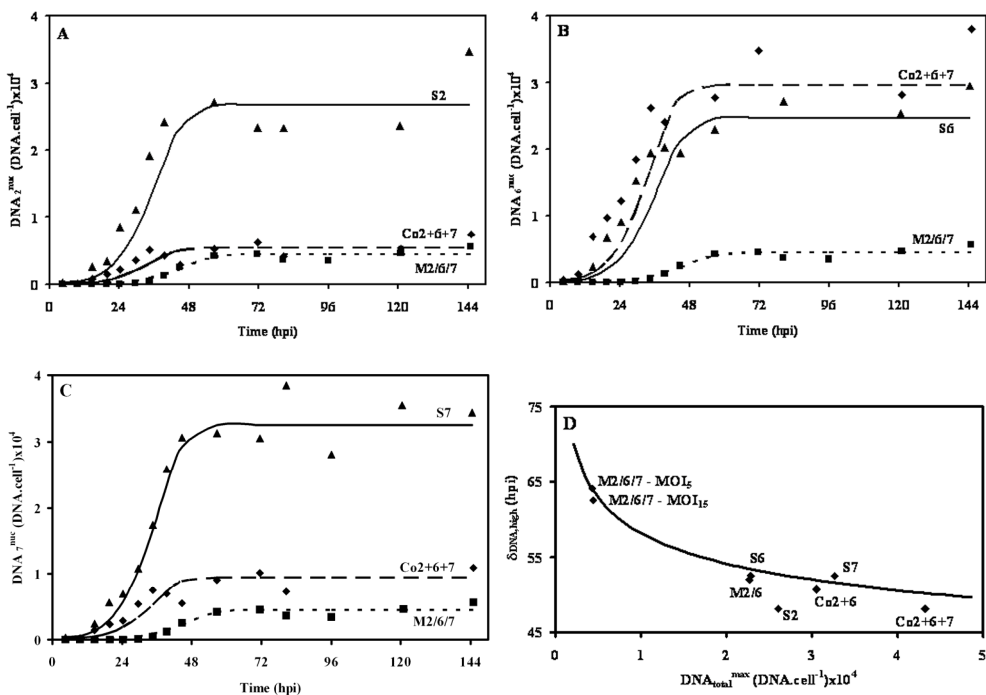


**Figure 28.** Dynamics of infected cell population. The lines and symbols denote model simulations and measurements respectively: experiment S2 (full line, —, open triangles,  $\Delta$ ), experiment S6 (dashed line, ---, open squares,  $\square$ ), experiment S7 (dotted line, ..., open circles,  $\circ$ ), experiment Co2+6+7 (dash and dot line, -.-.-, open diamonds,  $\diamond$ ) and experiment S2/6/7 (dash-dot-dot line, -.-.-.-, closed triangles,  $\blacktriangle$ ).

### 4.3. The dynamics of intracellular vDNA

Figure 29 shows vDNA measurements and model predictions for the experiments S2, S6, S7, Co2+6+7 and S2/6/7. The vDNA replication patterns are well described by first order replication kinetics, which is consistent with template-limited replication. The apparent vDNA replication rate is  $0.19 \pm 0.01 \text{ h}^{-1}$  and appears to be independent of the construct used. Indeed, in the monocistronic and multicistronic vectors used in this work,

the size of genes coding for the three viral proteins represents less than 5% of the baculovirus genome size [30]. Therefore the genome size factor is expected to have a minor effect on the replication kinetics. At least nine genes are known to be involved in vDNA replication [31]. The expression of these genes, including the *dnapol* and *dnahe1*, seems not to be a differentiation factor between the monocistronic and multicistronic vectors.



**Figure 29.** Dynamics of intracellular vDNA2 (A), vDNA6 (B) and vDNA7 (C). The lines and symbols denote model simulations and experimental data respectively: experiments S2, S6 and S7 (full line, —, closed triangles, ▲), experiment Co2+6+7 (dashed line, ---, closed diamonds, ◆) and experiment M2/6/7 (dotted line, ···, closed squares, ■). The evaluation of total intracellular vDNA effect on the time-schedule for viral replication (metabolic decay) was also assessed based on equations 5.17 to 5.19 (D). The full line (—) represents model simulation and closed diamonds (◆) experimental data.

The concentration of intracellular vDNA converges asymptotically to a maximum intracellular load (Figures 29A to 29C). Rosinski et al. (2002) suggested that the progressive decrease in vDNA replication rates could be

caused by a virus shift from DNA polymerase to its own dedicated viral RNA polymerase, which only acts on its unique promoter motifs. This hypothesis was supported by the observation that the maximum vDNA content coincided with the onset of very late gene expression, at approximately 20 hpi for a MOI of 20 virus.cell<sup>-1</sup>. In our system, the maximum vDNA levels are reached much later, between 48 hpi and 60 hpi, thus refuting the aforementioned hypothesis.

This suggests that the halt of vDNA replication is practically coincident in time, with few exceptions, with the halt of protein expression and also with the increase in cell death rate. Therefore, this behaviour is here interpreted as being related with the intrinsic metabolic decay due to infection. In our model, the effect of the metabolic decay was expressed by equations 5.17 and 5.19 as a function of intracellular vDNA. Figure 29D illustrates this dependency. The MOI plays here an important role. High MOIs result in faster vDNA replication and eventually in the reduction of the time interval for vDNA replication and protein expression. This effect has been previously observed in several studies [18; 20; 21; 32].

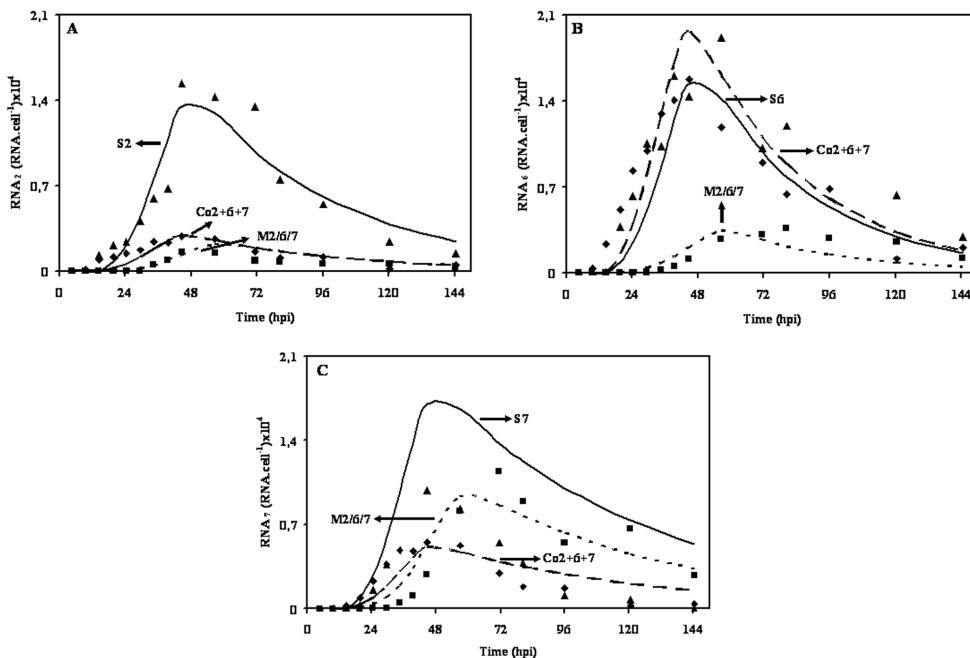
#### 4.4. Virus budding

In theory, the apparently constant vDNA concentration for  $t > \delta_D$  (hpi) could result from vDNA release into the extracellular medium through virus budding that would compensate for vDNA replication (virus budding is known to commence around 17-20 hpi [12; 33]). However, according to Rosinski et al. (2002), virus budding is considerably lower than the intracellular replication. This suggests that the virus makes much more vDNA than it is able to encapsidate and release. To confirm this hypothesis, extracellular vDNA was quantified for the co-infection experiment Co2+6+7. The maximum extracellular vDNA obtained was  $2.54 \times 10^9$  copies.ml<sup>-1</sup> (**DNA<sub>2</sub>+DNA<sub>6</sub>+DNA<sub>7</sub>**) at time  $t = 80.5$  hpi, which

represents 11% of the intracellular vDNA, thus confirming that virus budding is much lower than vDNA replication for the MOI ranges used.

#### 4.5. mRNA synthesis

Figure 30 shows mRNA measurements and model predictions for the experiments S2, S6, S7, Co2+6+7 and S2/6/7.



**Figure 30.** Intracellular dynamics of RNA2 (A), RNA6 (B) and RNA7 (C). The lines and symbols denote model simulations and experimental data respectively: experiments S2, S6 and S7 (full line, —, closed triangles, ▲), experiment Co2+6+7 (dashed line, ---, closed diamonds, ◆) and experiment M2/6/7 (dotted line, ...., closed squares, ■).

Typically, intracellular mRNA increases fast at the onset of the infection cycle up to a maximum level after which it starts decaying. The increase in mRNA concentration follows closely the increase in vDNA. The maximum marks the time instant when the mRNA transcription rate becomes lower than the corresponding degradation rate. The halt in mRNA synthesis is practically coincident in time with the halt in vDNA replication. In general,

the higher the intracellular vDNA levels the higher are the corresponding mRNA levels. The only exception to this scenario was verified for the gene coding for **vp<sub>7</sub>**, whose transcription is controlled by different promoters: *polh* in monocistronic vectors (Exp. 1 to 4) and *p10* in multicistronic vectors (Exp. 5).

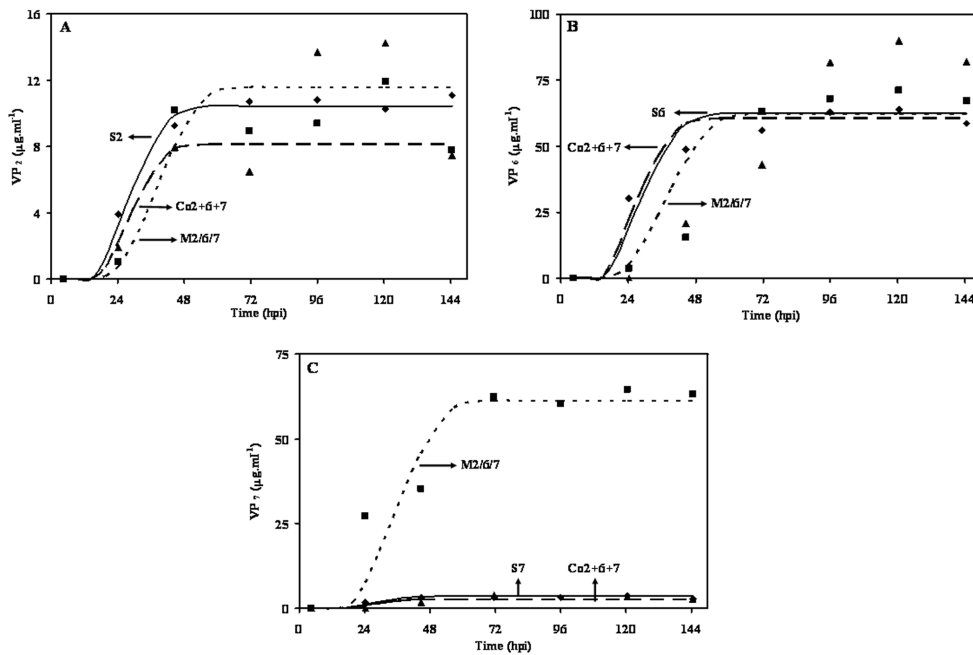
The synthesis of mRNA is well described by first order transcription kinetics. The characteristic transcriptional rates for the genes coding for **vp<sub>2</sub>**, **vp<sub>6</sub>** and **vp<sub>7</sub>** under the control of the *polh* promoter are  $0.12 \pm 0.01 \text{ h}^{-1}$ ,  $0.16 \pm 0.02 \text{ h}^{-1}$  and  $0.11 \pm 0.02 \text{ h}^{-1}$  respectively. These values fail to conform with the rule of the size of the genes: the smaller the gene the faster it is transcribed (the sizes of the genes coding for **vp<sub>2</sub>**, **vp<sub>6</sub>** and **vp<sub>7</sub>** are 2690, 1356 and 1062 bp respectively). Presuming a characteristic RNA polymerase elongation rate profile, the resulting transcriptional rates should be such that  $k_{\text{srna},7} > k_{\text{srna},6} > k_{\text{srna},2}$ . Since the aforementioned rates were consistent for S2, S6, S7 and Co2+6+7 experiments, a possible justification for this ranking is the use of different restriction sites for cloning. The transcription of *polh* and *p10* genes, starts within a baculovirus late promoter motif TAAG [34]. Modifications in the 5' untranslated regions (UTRs), between the TAAG motif and the ATG translational start codon, and in the 3' UTR of the baculovirus genome are known to have some influence in the translation and transcription levels [35-40].

It should be noticed that the previous transcription rates were the same for all the experiments except for the transcription of the gene coding for **vp<sub>7</sub>** (under the control of *p10* promoter) in M2/6/7 (Exp. 5), which was  $0.46 \pm 0.01 \text{ h}^{-1}$ . This exceptionally high  $k_{\text{srna},7}$  resulted in a higher **mRNA<sub>7</sub>** concentration when compared to Co2+6+7, even though the concentration of **vDNA<sub>7</sub>** was lower (Figure 29C vs 30C). The justification for this variation cannot be directly attributed to the strength of the promoter since the levels

of *polh* and *p10* driven expression are expected to be similar [41]. The interaction between *polh* and *p10* promoters are however known to influence the level of gene expression [42]. A relevant point is that, for the M2/6/7 experiment, where the transcription of the viral genes occurs from the same DNA template, the ranking of the transcriptional rates according to gene size is obeyed.

#### 4.6. Protein synthesis

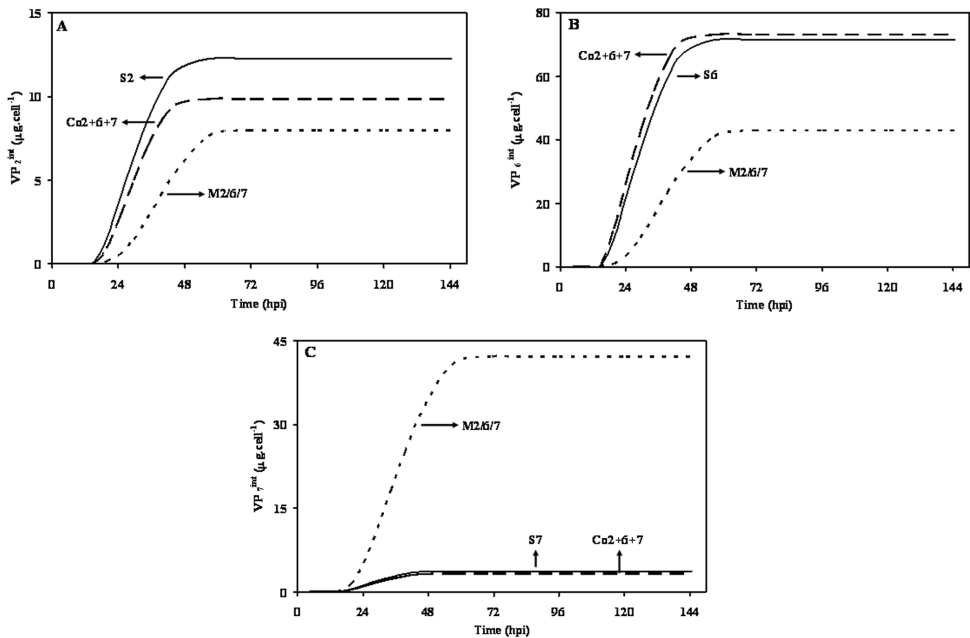
Figure 31 shows the dynamics of measured total (intracellular plus extracellular) protein,  $VP_i$  ( $\mu\text{g}\cdot\text{ml}^{-1}$ ), whereas Figure 32 shows model predictions of intracellular protein only.



**Figure 31.** Total (intracellular plus extracellular) protein dynamics of vp2 (A), vp6 (B) and vp7 (C). The lines and symbols denote model simulations and experimental data respectively: experiments S2, S6 and S7 (full line, —, closed triangles,  $\blacktriangle$ ), experiment Co2+6+7 (dashed line, ---, closed diamonds,  $\blacklozenge$ ) and experiment M2/6/7 (dotted line, ..., closed squares,  $\blacksquare$ ).



The dynamics of intracellular protein and vDNA show similar patterns. Intracellular protein concentration  $VP_j^{int}$  converges asymptotically to a maximum level that is highly correlated with the corresponding maximum vDNA level. Indeed, higher intracellular vDNA concentrations induce higher mRNA transcription levels and ultimately higher intracellular protein concentrations. This was verified, whenever the gene expression was under the control of the *polh* promoter. The unique exception was the gene coding for **vp<sub>7</sub>** in M2/6/7 (under the control of *p10* promoter), where significantly higher **mRNA<sub>7</sub>** concentrations were achieved from significantly lower **DNA<sub>7</sub>** templates (Figures 29C and 30C).



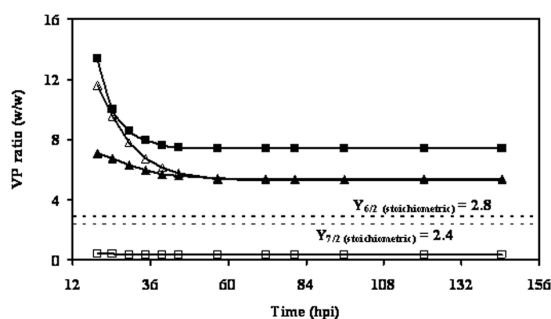
**Figure 32.** Intracellular vp2 (A), vp6 (B) and vp7 (C) concentrations. The lines represent model simulations: the full line (—) denotes S2, S6 and S7 strategies, the dashed line (---) strategy Co2+6+7 and the dotted line (....) strategy M2/6/7.

Protein synthesis seems to be reasonably well described by Michaelis-Menten kinetics on mRNA. The maximum protein synthesis rate varies considerably (see Table 17). The rate of protein synthesis may vary due to

codon usage and/or size of protein and also the time for maturation. The **vp<sub>7</sub>** formation takes more time, approximately 0.5 h [43] since it is a glycosylated protein.

#### 4.7. Analysis of protein dynamics

The following analysis will be focused on the dynamics of protein stoichiometric ratios and on their comparison with the composition of correctly assembled RLPs in the Co2+6+7 and M2/6/7 experiments. The relative mass composition of **vp<sub>6</sub>/vp<sub>2</sub>** and **vp<sub>7</sub>/vp<sub>2</sub>** in correctly assembled particles is  $Y_{6/2} = 2.8$  (w/w) and  $Y_{7/2} = 2.4$  (w/w) respectively. Figure 33 shows the dynamics of  $Y_{6/2}$  and  $Y_{7/2}$  for the M2/6/7 and Co2+6+7 strategies. It is clear that in the M2/6/7 experiment there is an excess of **vp<sub>7</sub>**, especially during the first 48 hpi. During this period, approximately 80% (w/w) of the final **vp<sub>7</sub>** was already synthesized (see Figure 31C). It is likely that the **vp<sub>7</sub>** outer layer can only be formed after the **vp<sub>2</sub>** and **vp<sub>6</sub>** layers are correctly assembled. Therefore, such high **vp<sub>7</sub>** levels at the beginning will probably favour the assembly of malformed particles. The experimental  $Y_{6/2}$  is systematically above the theoretical value; hence, it is likely that **vp<sub>2</sub>** is the limiting protein for the assembly of RLP.



**Figure 33.** Dynamics of proteins stoichiometric ratios. The dotted lines represent the relative mass composition of **vp<sub>6</sub>/vp<sub>2</sub>** ( $Y_{6/2} = 2.8$  (w/w)) and **vp<sub>7</sub>/vp<sub>2</sub>** ( $Y_{7/2} = 2.4$  (w/w)) in correctly assembled particles. Model simulations are denoted by symbols: closed squares (■) indicate  $Y_{6/2}$  whereas open squares (□) represents  $Y_{7/2}$  in Co2+6+7 strategy; closed triangles (▲) and open triangles (△) represents  $Y_{6/2}$  and  $Y_{7/2}$  respectively in M2/6/7 strategy.

Analyzing the Co2+6+7 strategy, experimental values of  $Y_{7/2}$  are lower than the theoretical value. Thus, this strategy is prone to produce particles composed of **vp**<sub>2</sub> and **vp**<sub>6</sub> layers (**vp**<sub>6</sub> is not limiting since  $Y_{6/2} > 2.8$  (w/w)), but with the outer **vp**<sub>7</sub> shell only partially formed or even completely absent. The volumetric productivity of M2/6/7 and Co2+6+7 strategies, after purification, reflects the observed strong under-expression of **vp**<sub>7</sub> in the Co2+6+7, leading to a lower RLP titer: 430 µg/ml in Co2+6+7 and 662 µg/ml in M2/6/7.

In the end, single-infection strategies appear to be advantageous over co-infection for RLP production. This suggests that process optimization can be achieved through the redesign of recombinant baculovirus vectors, namely at the promoters' level since different promoters may induce different levels of expression at more convenient times for correct particle formation. On the other hand, the co-infection strategy offers additional degrees of freedom for process optimization and has thus gained more attention in recent years. Theoretically, the time evolution of **vp**<sub>2</sub>:**vp**<sub>6</sub>:**vp**<sub>7</sub> stoichiometric ratio can be manipulated by the MOI of the respective baculovirus and, to a less extent, by scheduling the time of infection (TOI) of individual monocistronic baculoviruses. In addition, RLPs production can be significantly improved by controlling the cell concentration at infection (CCI), as long as specific productivities could be maintained.

## 5. CONCLUSIONS

In this work we proposed a deterministic mathematical model of the baculovirus/insect cell system for the production of viral proteins and RLPs. The model addresses the most critical steps of baculovirus internalization and trafficking, and gene expression mechanisms. The main objective of

this modelling study was to understand and quantify the effect on viral protein synthesis and RLP formation of different infection strategies, and evaluate the effect of different promoters. This modelling tool may allow redesigning better baculovirus vectors and is a starting point to optimize infection parameters such as MOI, TOI or CCI in co-infection experiments. The proposed model represents a good compromise between complexity and accuracy. A relatively small number of kinetic parameters are involved, which were estimated with high statistical confidence in most cases. From the analysis of the modelling results, the following more specific conclusions may be outlined:

- A characteristic apparent vDNA replication rate of  $0.19 \pm 0.01 \text{ h}^{-1}$  was obtained for monocistronic and multicistronic baculovirus constructs;
- Virus budding is lower than vDNA replication for  $\text{MOI} > 2 \text{ virus.cell}^{-1}$ ;
- The vDNA translation and transcription is well described by first-order and Michaelis-Menten kinetics, respectively. Reproducible kinetic parameters for the *polh* and *p10* promoters were obtained;
- The *p10* promoter in the M2/6/7 leads consistently to a much higher **vp<sub>7</sub>** expression than when *polh* controls **vp<sub>7</sub>** expression in the Co2+6+7 and S7 experiments;
- The time windows for effective vDNA, mRNA and protein synthesis are highly correlated. High MOIs tend to increase vDNA replication, but, at the same time, the metabolic burden is more severe, thereby decreasing the time window for effective protein expression;
- In terms of RLP synthesis, single-infection experiment leads to both **vp<sub>6</sub>** and **vp<sub>7</sub>** over-expression. However, the high **vp<sub>7</sub>** expression at the beginning of the infection cycle is undesirable. It is the expression of **vp<sub>2</sub>** that limits the assembly of RLPs;
- In co-infection experiments, **vp<sub>7</sub>** is largely under-expressed. Thus, probably particles of **vp<sub>2</sub>** and **vp<sub>6</sub>** with incomplete **vp<sub>7</sub>** shell are formed.

Optimized infection strategies are required to increase the expression of vp<sub>7</sub>.

## 6. ACKNOWLEDGEMENTS

The authors wish to thank Catarina Estêvão and Maria Candida Mellado for thoughtful discussions. The authors are also grateful to Dr Pedro Cruz for excellent advice, to Dr. John Aunins and Dr. Luis Maranga from Merck & Co., West Point, USA, and Prof. Hansjörg Hauser, GBF, Germany for valuable comments on the manuscript. This work was supported by European Commission project (QLRT-2001-01249) and the Portuguese Fundação para a Ciência e Tecnologia (POCTI/BIO/55975/2004 and SFRH/BD/21910/2005).

## 7. REFERENCES

1. Parashar, U.D., Holman, R.C., Cummings, K.C., Staggs, N.W., Curns, A.T., Zimmerman, C.M., Kaufman, S.F., Lewis, J.E., Vugia, D.J., Powell, K.E., Glass, R.I., **2000**. Trends in intussusception-associated hospitalizations and deaths among US infants. *Pediatrics*. 106 (6): 1413-21.
2. Heaton, P.M., Goveia, M.G., Miller, J.M., Offit, P., Clark, H.F., **2005**. Development of a pentavalent rotavirus vaccine against prevalent serotypes of rotavirus gastroenteritis. *J Infect Dis*. 192 Suppl 1 S17-21.
3. Bernstein, D.I., **2006**. Live attenuated human rotavirus vaccine, Rotarix. *Semin Pediatr Infect Dis*. 17 (4): 188-94.
4. Van der Wielen, M., Van Damme, P., **2008**. Pentavalent human-bovine (WC3) reassortant rotavirus vaccine in special populations: a review of data from the Rotavirus Efficacy and Safety Trial. *Eur J Clin Microbiol Infect Dis*. 27 (7): 495-501.
5. McCormack, P.L., Keam, S.J., Bernstein, D.I., Grimwood, K., Linhares, A.C., Madhi, S.A., Nakagomi, O., Vesikari, T., **2009**. Rotavirus vaccine RIX4414 (Rotarix): a review of its use in the prevention of rotavirus gastroenteritis. *Paediatr Drugs*. 11 (1): 75-88.
6. Labbe, M., Charpilienne, A., Crawford, S.E., Estes, M.K., Cohen, J., **1991**. Expression of rotavirus VP2 produces empty corelike particles. *J Virol*. 65 (6): 2946-52.
7. Prasad, B.V., Wang, G.J., Clerx, J.P., Chiu, W., **1988**. Three-dimensional structure of rotavirus. *J Mol Biol*. 199 (2): 269-75.
8. Crawford, S.E., Labbe, M., Cohen, J., Burroughs, M.H., Zhou, Y.J., Estes, M.K., **1994**. Characterization of virus-like particles produced by the expression of rotavirus capsid proteins in insect cells. *J Virol*. 68 (9): 5945-5952.
9. Vieira, H., Estêvão, C., Roldão, A., Peixoto, C., Sousa, M., Cruz, P.E., Carrondo, M.J.T., Alves, P.M., **2005**. Triple layered rotavirus VLP assembly: kinetics of vector replication, mRNA stability and recombinant protein production. *J Biotechnol.* 120 (1): 72-82.

10. Volkman, L.E., Goldsmith, P.A., **1985**. Mechanism of neutralization of budded *Autographa Californica* Nuclear Polyhedrosis Virus by a monoclonal antibody: Inhibition of entry by adsorptive endocytosis. *Virology*. 143 185-195.
11. Possee, R.D., King, L.A., **1992**. *The Baculovirus Expression System: A laboratory guide*, London.
12. Power, J.F., Nielsen, L.K., **1996**. Modelling baculovirus infection of insect cells in culture. *Cytotechnology*. 20 (1-3): 209-219.
13. Dee, K.U., Shuler, M.L., **1997**. A mathematical model of the trafficking of acid-dependent enveloped viruses: Application to the binding, uptake, and nuclear accumulation of baculovirus. *Biotechnol Bioeng*. 54 (5): 468-490.
14. Dee, K.U., Hammer, D.A., Shuler, M.L., **1995**. A model of the binding, entry, uncoating, and RNA synthesis of Semliki Forest virus in baby hamster kidney (BHK-21) cells. *Biotechnol Bioeng*. 46 (5): 485-496.
15. Wickham, T.J., Granados, R.R., Wood, H.A., Hammer, D.A., Shuler, M.L., **1990**. General analysis of receptor-mediated viral attachment to cell surfaces. *Biophys J*. 58 (6): 1501-1516.
16. Wickham, T.J., Shuler, M.L., Hammer, D.A., Granados, R.R., Wood, H.A., **1992**. Equilibrium and kinetic analysis of *Autographa californica* nuclear polyhedrosis virus attachment to different insect cell lines. *J Gen Virol*. 73 (Pt 12): 3185-3194.
17. Dalal, N.G., Bentley, W.E., **1999**. Mathematical characterization of insect cell (Sf-9) death in suspended culture. *Biotechnol Lett*. 21 325-329.
18. Hu, Y.-C., Bentley, W.E., **2000**. A kinetic and statistical-thermodynamic model for baculovirus infection and virus-like particle assembly in suspended insect cells. *Chem Eng Sci*. 55 3991-4008.
19. Dee, K.U., Shuler, M.L., **1997**. Optimization of an assay for baculovirus titer and design of regimens for the synchronous infection of insect cells. *Biotechnol Prog*. 13 (1): 14-24.
20. Hu, Y.C., Bentley, W.E., **2001**. Effect of MOI ratio on the composition and yield of chimeric infectious bursal disease virus-like particles by baculovirus co-infection: deterministic predictions and experimental results. *Biotechniques*. 75 (1): 104-119.
21. Licari, P., Bailey, J.E., **1992**. Modeling the population dynamics of baculovirus-infected insect cells: Optimizing infection strategies for enhanced recombinant protein yields. *Biotechnol Bioeng*. 39 (4): 432-441.
22. Palomares, L.A., Lopez, S., Ramirez, O.T., **2002**. Strategies for manipulating the relative concentration of recombinant rotavirus structural proteins during simultaneous production by insect cells. *Biotechnol Bioeng*. 78 (6): 635-644.
23. Mena, J.A., Ramirez, O.T., Palomares, L.A., **2005**. Intracellular distribution of rotavirus structural proteins and virus-like particles expressed in the insect cell-baculovirus system. *J Biotechnol*.
24. Tosser, G., Labbe, M., Bremont, M., Cohen, J., **1992**. Expression of the major capsid protein VP6 of group C rotavirus and synthesis of chimeric single-shelled particles by using recombinant baculoviruses. *J Virol*. 66 (10): 5825-31.
25. Peixoto, C., Sousa, M.F., Silva, A.C., Carrondo, M.J., Alves, P.M., **2007**. Downstream processing of triple layered rotavirus like particles. *J Biotechnol*. 127 (3): 452-61.
26. Mellado, M.C., Franco, C., Coelho, A., Alves, P.M., Simplicio, A.L., **2008**. Sodium dodecyl sulfate-capillary gel electrophoresis analysis of rotavirus-like particles. *J Chromatogr A*. 1192 (1): 166-72.
27. Roldão, A., Oliveira, R., Carrondo, M.J.T., Alves, P.M., **2009**. Error Assessment in Recombinant Baculovirus Titration: Evaluation of Different Methods. *J Virol Methods*.
28. Roldão, A., Vieira, H.L., Charpilienne, A., Poncet, D., Roy, P., Carrondo, M.J., Alves, P.M., Oliveira, R., **2007**. Modeling rotavirus-like particles production in a baculovirus expression vector system: Infection kinetics, baculovirus DNA replication, mRNA synthesis and protein production. *J Biotechnol*. 128 (4): 875-894.
29. Licari, P., Bailey, J.E., **1991**. Factors influencing recombinant protein yields in an insect cell-baculovirus expression system: Multiplicity of infection and intracellular protein degradation. *Biotechnol Bioeng*. 37 (3): 238-246.
30. Summers, M.D., Smith, G.E., **1978**. Baculovirus structural polypeptides. *Virology*. 84 (2): 390-402.
31. Lu, A., Miller, L.K., **1995**. The roles of eighteen baculovirus late expression factor genes in transcription and DNA replication. *J Virol*. 69 (2): 975-82.
32. Rosinski, M., Reid, S., Nielsen, L.K., **2002**. Kinetics of baculovirus replication and release using real-time quantitative polymerase chain reaction. *Biotechnol Bioeng*. 77 (4): 476-480.
33. Wong, K. **1997**. Optimization of recombinant protein production using the baculovirus expression system: Characterization and modeling of infectious events., Department of Chemical Engineering, Brisbane, Australia.
34. Blissard, G.W., Rohrmann, G.F., **1990**. Baculovirus diversity and molecular biology. *Annu Rev Entomol*. 35 127-155.
35. van Oers, M.M., Vlak, J.M., Voorma, H.O., Thomas, A.A., **1999**. Role of the 3' untranslated region of baculovirus p10 mRNA in high-level expression of foreign genes. *J Gen Virol*. 80 (Pt 8): 2253-2262.

36. Qin, J.C., Liu, A.F., Weaver, R.F., **1989**. Studies on the control region of the p10 gene of the *Autographa californica* nuclear polyhedrosis virus. *J Gen Virol.* 70 (Pt 5): 1273-1279.
37. Ooi, B.G., Rankin, C., Miller, L.K., **1989**. Downstream sequences augment transcription from the essential initiation site of a baculovirus polyhedrin gene. *J Mol Biol.* 210 (4): 721-736.
38. Weyer, U., Possee, R.D., **1988**. Functional analysis of the p10 gene 5' leader sequence of the *Autographa californica* nuclear polyhedrosis virus. *Nucleic Acids Res.* 16 (9): 3635-3653.
39. Luckow, V.A., Summers, M.D., **1988**. Signals important for high-level expression of foreign genes in *Autographa californica* nuclear polyhedrosis virus expression vectors. *Virology.* 167 (1): 56-71.
40. Possee, R.D., Howard, S.C., **1987**. Analysis of the polyhedrin gene promoter of the *Autographa californica* nuclear polyhedrosis virus. *Nucleic Acids Res.* 15 (24): 10233-10248.
41. van Oers, M.M., Malarme, D., Jore, J.M., Vlak, J.M., **1992**. Expression of the *Autographa californica* nuclear polyhedrosis virus p10 gene: effect of polyhedrin gene expression. *Arch Virol.* 123 (1-2): 1-11.
42. Chaabihi, H., Ogliastro, M.H., Martin, M., Giraud, C., Devauchelle, G., Cerutti, M., **1993**. Competition between baculovirus polyhedrin and p10 gene expression during infection of insect cells. *J Virol.* 67 (5): 2664-2671.
43. Lodish, H., Baltimore, D., Darnell, J., 1986. *Molecular Cell Biology*. Scientific American Books, Inc., New York, USA.





# Chapter 6

Optimization of RLP production using  
the stochastic model

## Summary

In this study, the theoretical effect of MOI and promoter strength on RLP production at single and co-infection strategies was investigated using the stochastic model developed in Chapter 3.

RLP production can be optimized in single-infection strategies by manipulation of the MOI (lower MOIs) and/or by redesigning baculovirus vectors, in which the promoters controlling viral protein expression can be selected according to their strength (lowering the strength of the promoter controlling **vp7** expression in relation to the other two promoters). At co-infection strategies, high MOIs maximize RLP production; the redesign of baculovirus vectors configures a somewhat modest improvement potential.

At low MOIs, single-infection strategies offer significant advantages over co-infection experiments, namely: 1)  $n_{i,267}$  can reach up to  $3 \times 10^6$  cell.ml<sup>-1</sup>, contrasting with the  $0.6 \times 10^6$  cell.ml<sup>-1</sup> at co-infection; 2) higher RLP concentrations: 30-60 mg.L<sup>-1</sup> against the 6-12 mg.L<sup>-1</sup> at co-infection; 3) assembly efficiencies are higher: 72% against the 20-40% at co-infection. At high MOIs, the differences in model outputs for co- and single-infection strategies are meaningless.

The findings described throughout this chapter demonstrate that the redesign of baculovirus vectors, in which the promoters controlling viral protein expression can be selected according to their strength, and the manipulation of the MOI have great potential for process optimization. The practical implementation of these concepts may not be easy since it depends on the available resources, namely the type of recombinant baculoviruses (tricistronic or monocistronic), and on the knowledge of molecular biology techniques, essential for the construction of novel recombinant baculoviruses.

## CONTENTS

<b>1. Introduction .....</b>	<b>188</b>
<b>2. The mathematical model .....</b>	<b>189</b>
<b>3. Sensitivity analysis of model parameters .....</b>	<b>191</b>
<b>4. Optimization of RLPs production .....</b>	<b>200</b>
4.1. Single-infection strategies .....	200
4.1.1. The effect of promoter strength .....	202
4.2. Co-infection strategies .....	206
4.2.1. Are traditional co-infection strategies suboptimal? .....	210
4.2.2. Promoter effect at co-infection strategies .....	211
4.3. What to choose: single-infection or co-infection? .....	213
<b>5. Conclusions and future perspectives .....</b>	<b>214</b>
<b>6. Acknowledgements .....</b>	<b>217</b>
<b>7. References .....</b>	<b>217</b>

## 1. INTRODUCTION

One of the most promising immunization strategies against rotavirus disease consists on a subunit vaccine based on virus-like particles produced in the baculovirus/insect cell system - rotavirus-like particles (RLP) [1-4]. Yet, production of this vaccine candidate has been hindered by the complexity of the RLP production process, characterized by low volumetric productivities and high amounts of unassembled proteins (up to 88% of the total mass of proteins expressed) [5-12].

Increased productivity and assembly efficiency can be achieved through the redesign of baculovirus vectors and through the infection strategy. Both factors are interlinked and cannot be analyzed separately. Two possible strategies can be pursued: 1. co-infection, where cells are infected with three monocistronic baculoviruses enclosing the genes coding for rotavirus proteins **vp2**, **vp6** or **vp7**; 2. single-infection strategies, where cells are infected with tricistronic baculoviruses enclosing the genes coding for the three viral proteins simultaneously [12]. In single-infection, the control of protein expression is achieved at the genetic level, while in the co-infection strategy the same is attempted at the process level.

Co-infection strategies allow to control the time and quantity of genes delivered into the insect cells. This strategy provides in theory some degree of control of relative protein expression titers, which is a key factor for assembly efficiency as demonstrated in Chapter 4. It is likely that productivities may be highly improved through manipulation of key process parameters such as the multiplicity of infection (MOI).

In single-infection strategies, different promoters may induce different expression levels at more convenient times; hence, the optimization of RLP titer and assembly efficiency can be achieved via baculovirus vector

redesign. In theory, using promoters with different strengths, it would be possible to make **vp2**, **vp6** and **vp7** expression levels coincide with the stoichiometric composition of the RLP, the optimal condition for particle assembly (Chapter 5).

This work aims at evaluating the effect of MOI (low and high) and promoter strength on RLP production for both scenarios of single and co-infection strategies using the stochastic model developed in Chapter 3 [13].

## 2. THE MATHEMATICAL MODEL

In this study, the stochastic model described in Chapter 3 was extended for simultaneous expression of the three viral proteins that compose RLP. The complete set of equations is presented in Table 18. Note that index  $k=2,6$  and 7 refer to **vp2**, **vp6** and **vp7** respectively, while in Chapter 2 only **vp2** was studied. Moreover, a stoichiometric assembly equation is included (last equation in Table 18). The assumptions behind this model were discussed previously in Chapter 3. The most critical of these, are the following:

- nutrients are in excess, thus, not limiting cellular growth;
- the sizes of the monocistronic and tricistronic baculovirus constructs are similar; thus, vDNA replication kinetics is assumed to be the same for the different constructs;
- a characteristic host ribosomal elongation rate of  $5.5 \times 10^6 \text{ bp.rna}^{-1}.\text{h}^{-1}$  and mRNA transcription rate of  $325 \text{ bp.h}^{-1}$  were assumed (estimated in Chapter 3). The main distinguishing factor in mRNA transcription rate and viral protein translation rate of **vp2**, **vp6** and **vp7** is the size of the underlying gene [13];

**Table 18.** Mathematical model equations.

Cellular growth	Uninfected cell population	$\frac{dn_u}{dt} = (\mu_{\max} - k_d)n_u$
	Infected cell population	$\frac{dn_i}{dt} = k_d n_i$
	Cell death	$k_d = \begin{cases} k_{d,int} + \Delta k_d \log_{10} \left( \sum_k dna_k \right), & t > TOI + \delta^* \\ k_{d,int}, & \text{otherwise} \end{cases}$
Baculovirus infection	Infection rate	$r_{inf,k} dt = \begin{cases} X_k, & t \leq TOI + \delta_{re-inf} \\ 0, & t > TOI + \delta_{re-inf} \end{cases}$
	Probability of infection	$P_{X_k=nk} = \frac{\exp(-k_a v_k(t) dt) (k_a v_k(t) dt)^{nk}}{nk!}$
	Extracellular virus	$\frac{dv_k}{dt} = -r_{inf,k} n_u + r_{bud,k} n_i$
vDNA replication	Intracellular vDNA	$\frac{d dna_k}{dt} = \eta_r r_{inf,k} + r_{dnarep,k} - r_{bud,k}, \quad dna_k(0) = 0$
	Replication rate	$r_{dnarep,k} = \begin{cases} k_{RDNA} dna_k \left( 1 - \frac{\delta^* - (t - TOI)}{\delta^* - \delta_{DNA,low}} \right), & \delta_{DNA,low} \leq t \leq \delta^* \\ 0, & \text{otherwise} \end{cases}$
	Halt in vDNA replication	$\delta^* = \frac{\delta_{10}}{\log_{10} \left( \sum_k dna_k \right)}$
Progeny viriogenesis	Virus budding	$r_{bud,k} = \begin{cases} k_{bud} dna_k, & \delta_{bud,low} \leq t - TOI \leq \delta_{bud,high} \\ 0, & \text{otherwise} \end{cases}$
mRNA synthesis	Intracellular RNA	$\frac{d rna_k}{dt} = r_{transc,k} - k_{DRNA,k} rna_k, \quad rna_k(0) = 0$
	Transcription rate	$r_{transc,k} = \begin{cases} k_{rna,elong} \frac{dna_k}{S_k} \left( 1 - \frac{\delta^* - (t - TOI)}{\delta^* - \delta_{polh,low}} \right), & \delta_{polh,low} \leq t - TOI \leq \delta^* \\ 0, & \text{otherwise} \end{cases}$
Protein synthesis	Intracellular protein	$\frac{dvp_k}{dt} = r_{transl,k}, \quad vp_k(0) = 0$
	Synthesis rate	$r_{transl,k} = \begin{cases} k_{rib,elong} \frac{rna_k}{S_k} mp_k \left( 1 - \frac{\delta^* - (t - TOI)}{\delta^* - \delta_{polh,low}} \right) i_{transl}, & \delta_{polh,low} \leq t - TOI \leq \delta^* \\ 0, & \text{otherwise} \end{cases}$
	Translation capacity	$i_{transl} = \begin{cases} 1, & \sum_k S_k rna_k \leq L_{transl,max} \\ \frac{L_{transl,max}}{\sum_k S_k rna_k}, & \sum_k S_k rna_k > L_{transl,max} \end{cases}$
RLP production	Limiting protein	$RLP = \min(vp2 \times [1 + Y_{6/2} + Y_{7/2}], vp6 \times [1 + 1/Y_{6/2} + Y_{7/2}/Y_{6/2}], vp7 \times [1 + 1/Y_{7/2} + Y_{6/2}/Y_{7/2}])$

$n_u$  – concentration of uninfected cells (cell.ml<sup>-1</sup>);  $\mu_{\max}$  – specific cell growth rate (0.03 h<sup>-1</sup>) [14];  $k_d$  – cell death rate (h<sup>-1</sup>);  $n_i$  – concentration of infected cells (cell.ml<sup>-1</sup>);  $k_{d,int}$  – intrinsic cell death rate of Sf-9 cells (0.0008 h<sup>-1</sup>) [10, 15, 16];  $\Delta k_d$  – increase in cell death rate due to infection (0.0017 ± 0.0003 h<sup>-1</sup>) [10]; **TOI** – time of infection (h);  $\delta^*$  – time required for the infection to accelerate host cell death (hpi); **dna<sub>k</sub>** – intracellular vDNA concentration delivered by virus **k**; **k** – denotes the foreign gene;  $r_{inf,k}$  – rate of virus **k** entry into Sf-9 cells (dna.cell<sup>-1</sup>h<sup>-1</sup>); **X<sub>k</sub>** – random variable denoting the number of viruses **k** per unit cell (dna.cell<sup>-1</sup>) which penetrate Sf-9 cells in the time interval [t, t+dt];  $\delta_{re-inf}$  – time period over which cells continue to be re-infected (10 hpi) [16]; **k<sub>a</sub>** – first order attachment constant (7.8×10<sup>-8</sup> ml.cell<sup>-1</sup>.h<sup>-1</sup>) [17]; **v<sub>k</sub>** –

concentration of extracellular virus  $k$  ( $\text{dna.ml}^{-1}$ );  $r_{\text{bud},k}$  - virus budding rate ( $\text{dna.cell}^{-1}.\text{h}^{-1}$ );  $\eta_k$  - trafficking efficiency (0.5) [18];  $r_{\text{dnarep},k}$  - vDNA replication rate ( $\text{dna.cell}^{-1}.\text{h}^{-1}$ );  $k_{\text{RDNA}}$  - first-order vDNA replication constant ( $0.49 \pm 0.03 \text{ h}^{-1}$ );  $\delta_{\text{DNA},\text{low}}$  - time for the onset of vDNA replication (6 hpi) [19, 20];  $\delta^*$  - time at which vDNA replication stops (hpi);  $\delta_{10}$  - time instant correspondent to 10 intracellular vDNA copies ( $233 \pm 7 \text{ hpi}$ ) [10];  $\delta_{\text{bud},\text{low}}$  and  $\delta_{\text{bud},\text{high}}$  - time of onset and halt of virus budding (16 and 52 hpi) [16, 21];  $k_{\text{bud}}$  - first-order virus budding constant ( $0.07 \pm 0.02 \text{ h}^{-1}$ );  $\text{rna}_k$  - intracellular concentration of mRNA coding for protein  $k$  ( $\text{ma.cell}^{-1}$ );  $r_{\text{transc},k}$  - transcription rate ( $\text{rna.cell}^{-1}.\text{h}^{-1}$ );  $k_{\text{DRNA},k}$  - first-order degradation rate ( $0.019 \pm 0.004 \text{ h}^{-1}$ ) [10];  $k_{\text{ma},\text{elong}}$  - RNA polymerase elongation rate ( $325 \pm 69 \text{ bp.h}^{-1}$ );  $S_k$  - size of gene  $k$  (2690 bp) [10];  $\delta_{\text{polh},\text{low}}$  - time post infection for the onset of polh promoter transcription (15 hpi) [10, 22];  $\text{vp}_k$  - concentration of intracellular viral protein  $k$  ( $\mu\text{g.cell}^{-1}$ );  $r_{\text{transl},k}$  - protein synthesis rate ( $\mu\text{g.cell}^{-1}.\text{h}^{-1}$ );  $k_{\text{rib},\text{elong}}$  - the ribosomal elongation rate ( $5.5 \pm 0.6 \times 10^6 \text{ bp.ma}^{-1}.\text{h}^{-1}$ );  $\text{mp}_k$  - molecular weight of  $\text{vp}_k$  ( $1.69 \times 10^{-13} \mu\text{g}$ ) [23];  $l_{\text{transl}}$  - adimensional parameter to account for the hypothesis of a maximum translation capacity;  $L_{\text{transl},\text{max}}$  - maximum translation capacity ( $1.1 \pm 0.3 \times 10^7 \text{ ma.bp.cell}^{-1}$ );  $Y_{6/2}$  - stoichiometric ratio of  $\text{vp6}/\text{vp2}$  in correctly assembled RLP;  $Y_{7/2}$  - stoichiometric ratio of  $\text{vp7}/\text{vp2}$  in correctly assembled RLP.

- there is a maximum translation capacity constraint [13] and, as such, **vp2**, **vp6** and **vp7** translation compete for host translation machinery. This is the factor that distinguishes the expression of a single protein from multiple proteins;
- virus budding is considered as it plays a key role at low MOIs [13];
- for simplicity and computational reasons, intracellular and total RLP production were calculated using the irreversible model (see Chapter 4), i.e. assembly efficiency is determined by the limiting protein and RLP stoichiometry (last equation Table 18).

Despite the obvious advantages in terms of computational time, this latter assumption is debatable. As shown in Chapter 4 (Figure 25), the irreversible model predicts higher efficiencies than the thermodynamic assembly model. However, as the  $\Delta G^0_{\text{RLP}}$  decreases, i.e. physicochemical conditions are optimal for RLP assembly, the equilibrium model converges to the irreversible model. As such, the present analysis assumes that the host provides the ideal thermodynamic conditions for particle assembly.

### 3. SENSITIVITY ANALYSIS OF MODEL PARAMETERS

Model parameters sensitivity analysis was performed for the case of single protein expression, namely **vp2**, at a constant MOI of  $0.1 \text{ virus.cell}^{-1}$ . Note

that this MOI was determined to be the optimal MOI for **vp2** production in Chapter 3. The effect of 5%, 10%, 25% and 50% model parameters perturbations, either negative or positive, upon cellular growth, progeny virus synthesis, intracellular kinetics and protein expression was determined (Figures 34 to 37).

Negative and positive perturbations in model parameters affect differently the output of the model. For example, a 50% increase in the value of  $\mu_{\max}$  (specific cell growth rate) induces a higher variation in the concentration of uninfected cells ( $n_u$ ) when compared to 50% decrease; the opposite occurs for  $k_a$  (first order attachment constant) and  $k_{\text{RDNA}}$  (first-order vDNA replication constant) (Figure 34A). This is also observed for  $k_{\text{RDNA}}$  in vDNA replication (Figure 36A) and for  $S_2$  (size of gene coding for **vp2**) in total protein concentration (Figure 37).

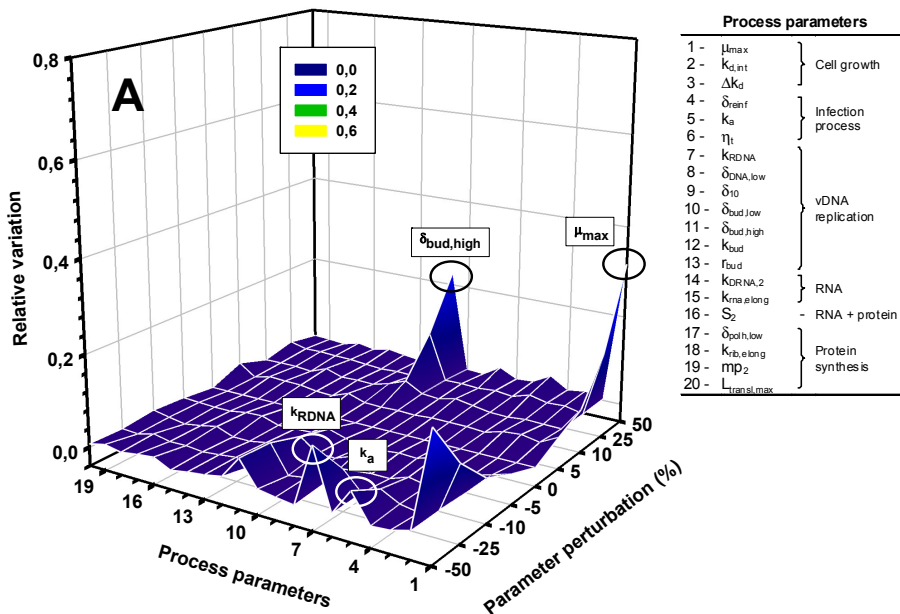
Figures 34A to 34C show that the different cell populations are not significantly affected by small perturbations in model parameters. Severe deviations are only observed when high perturbations are imposed. The parameters contributing most for such variation and that are common to every population analysed (uninfected, infected and death cell population) are the  $\mu_{\max}$ ,  $k_a$ ,  $k_{\text{RDNA}}$  and  $\delta_{\text{bud,high}}$  (time of halt of virus budding).

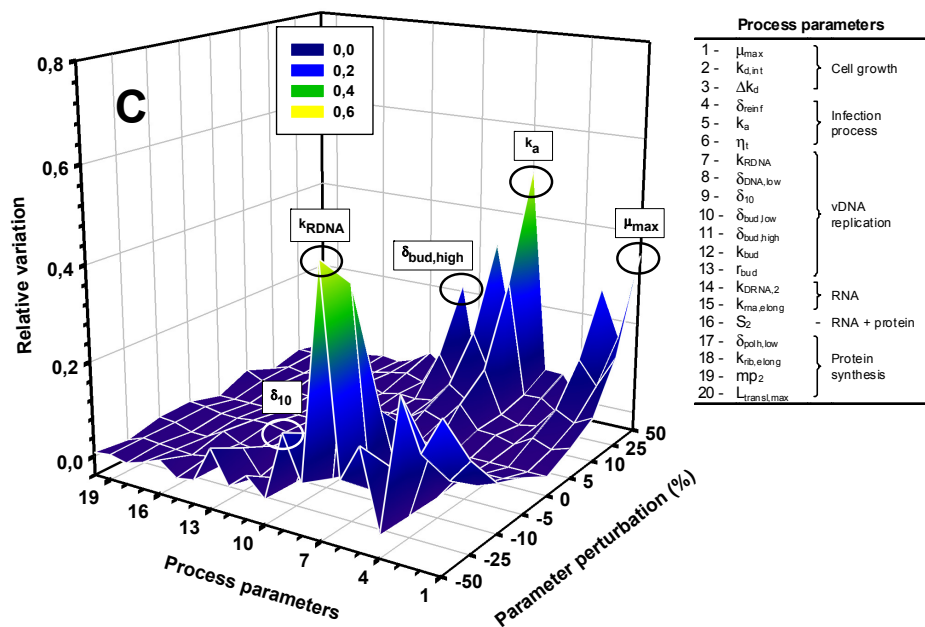
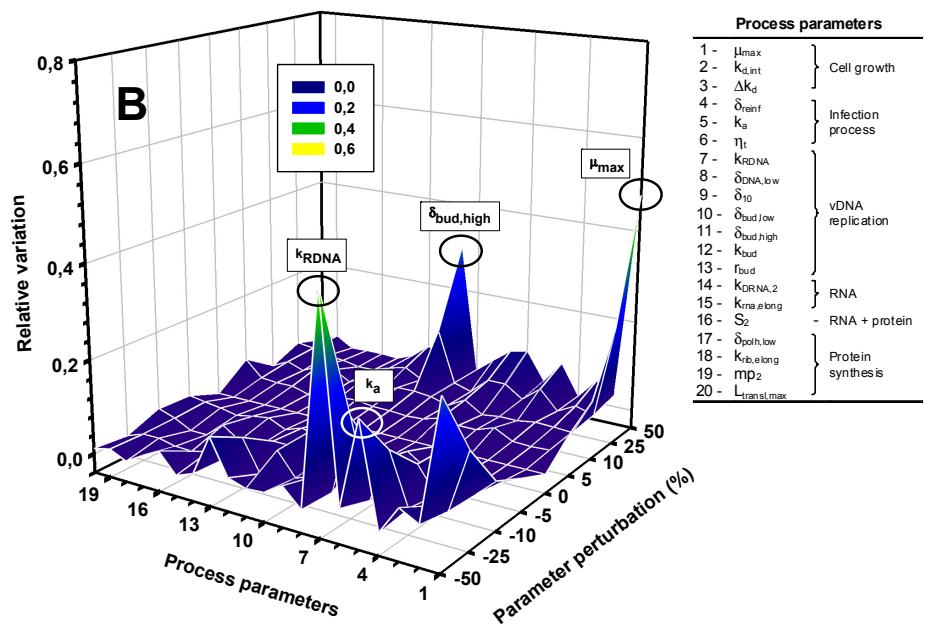
The first-order differential equation for cell growth clearly defines the correlation between uninfected cells concentration and  $\mu_{\max}$  (see Table 18). Since this relationship is exponentially-based, a small perturbation in  $\mu_{\max}$ , either positive or negative, is expected to induce a significant variation in the concentration of uninfected cells. This effect further propagates to the remaining cell populations analysed. Thus, the higher the perturbation in  $\mu_{\max}$ , the higher is the variation in the concentration of infected and death cells.



The variations observed in all cell populations upon perturbations in  $k_a$  can be interpreted in two different ways. One can speculate of the negative impact that  $k_a$  decrease has on intracellular vDNA concentration and progeny virus synthesis (see equations in Table 18), with its subsequent influence on second round of infection, characteristic of the MOI used ( $\text{MOI} = 0.1 \text{ virus.cell}^{-1}$  – asynchronous infection). However, the concentration levels of  $\text{dna}_2$  (Figure 36A) and  $v_2$  (Figure 35) are not significantly affected. This suggests that  $k_a$  acts directly on the concentration of infected cells; the lower the attachment rate of baculovirus to infected cells, the less efficient is the infection process. Uninfected and death cell populations are also affected by this parameter, although indirectly via the aforementioned variations in the concentration of infected cells.

**Figure 34.** Sensitivity analysis of uninfected,  $n_u$  (A,  $\text{cell.ml}^{-1}$ ), infected,  $n_{\text{inf}}$  (B,  $\text{cell.ml}^{-1}$ ) and death cells concentration,  $n_{\text{death}}$  (C,  $\text{cell.ml}^{-1}$ ) with respect to model parameter perturbations (%) (0-0.8 - relative variation of model outputs upon parameter perturbation with respect to model outputs prior to parameter perturbation).



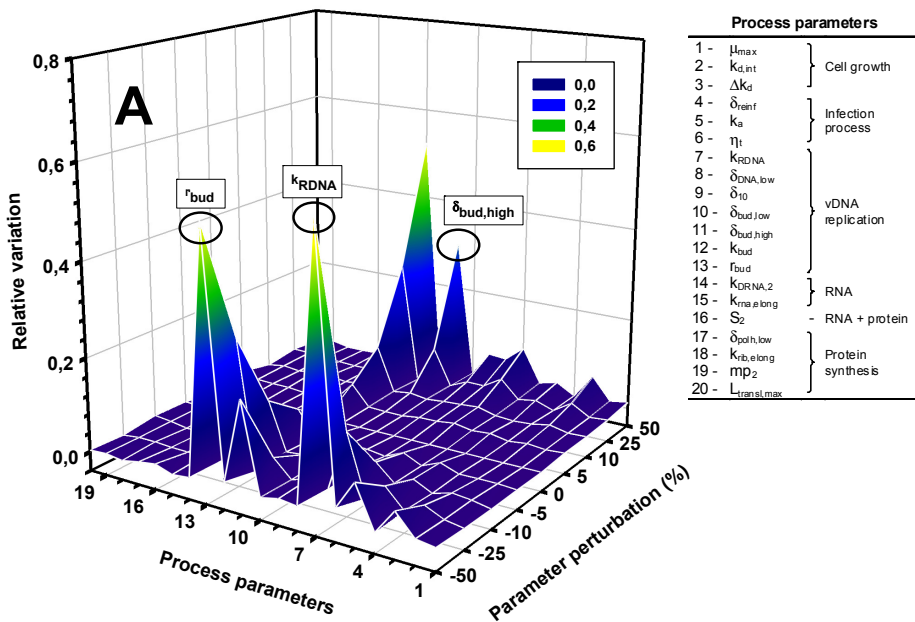


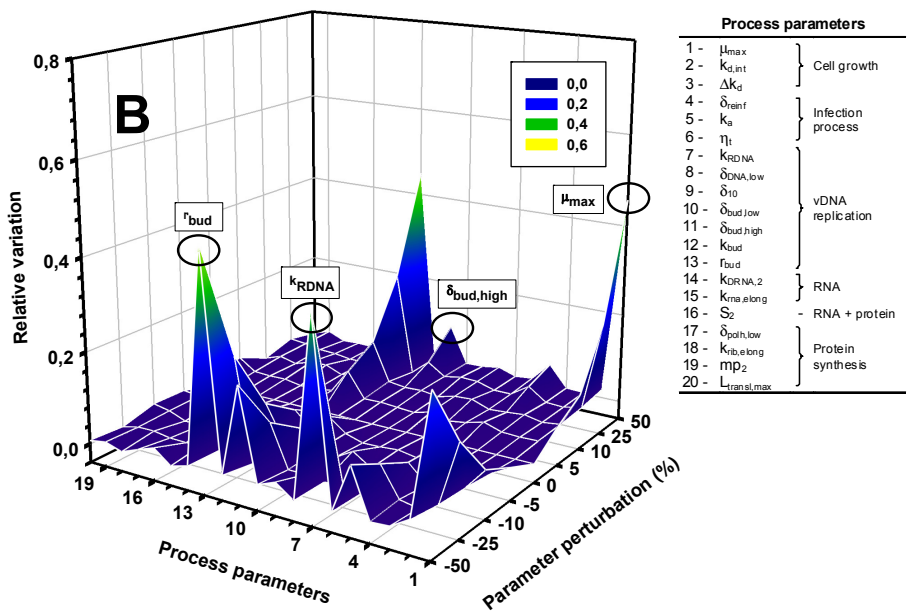
Variations in  $k_{RDNA}$  impact directly on the concentration of intracellular vDNA (Figure 36A) and indirectly on  $\Delta k_d$  (increase in cell death rate due to infection) and  $\delta^*$  (time required for the infection to accelerate host cell

death). Thus, the concentration of uninfected, infected and death cells are expected to be significantly affected.

The kinetics of progeny virus synthesis is highly sensitive to parameters such as the  $\delta_{\text{bud,high}}$  (time of halt of virus budding) and  $r_{\text{bud}}$  (virus budding rate) (Figure 35). In addition, changes in the specific cell growth rate,  $\mu_{\text{max}}$ , induce significant variations in total virus concentration whereas practically no effect is observed for intracellular virus production. This result is rational because the specific cell growth rate does not affect either the infection process or the replication machinery (intracellular production is maintained). It only impacts on the final concentration of infected cells, thus indirectly affecting the total concentration of viruses.

**Figure 35.** Sensitivity analysis of intracellular (A, virus.cell<sup>-1</sup>) and total (B, virus.ml<sup>-1</sup>) virus concentration ( $v_2$ ) with respect to model parameter perturbations (%) (0-0.8 - relative variation of model outputs upon parameter perturbation with respect to model outputs prior to parameter perturbation).

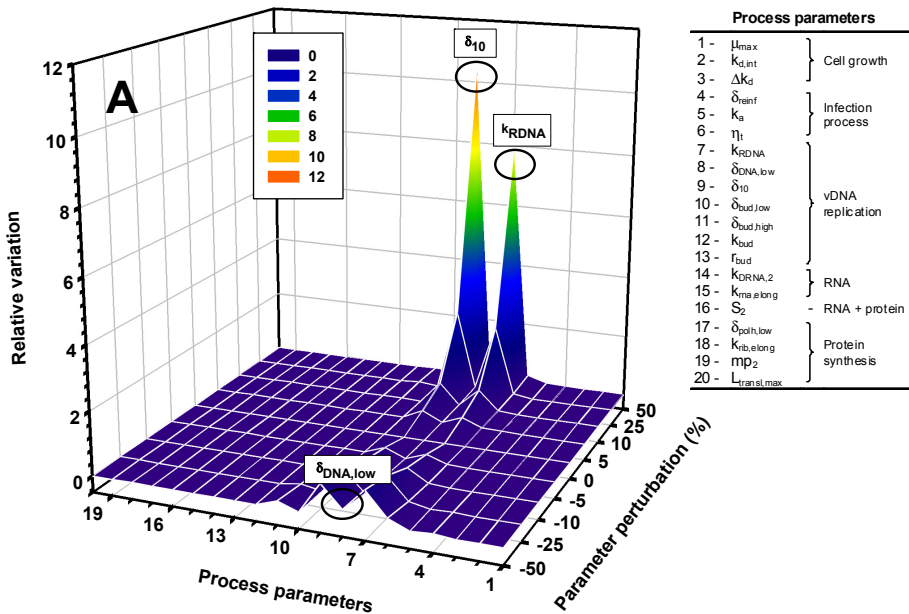


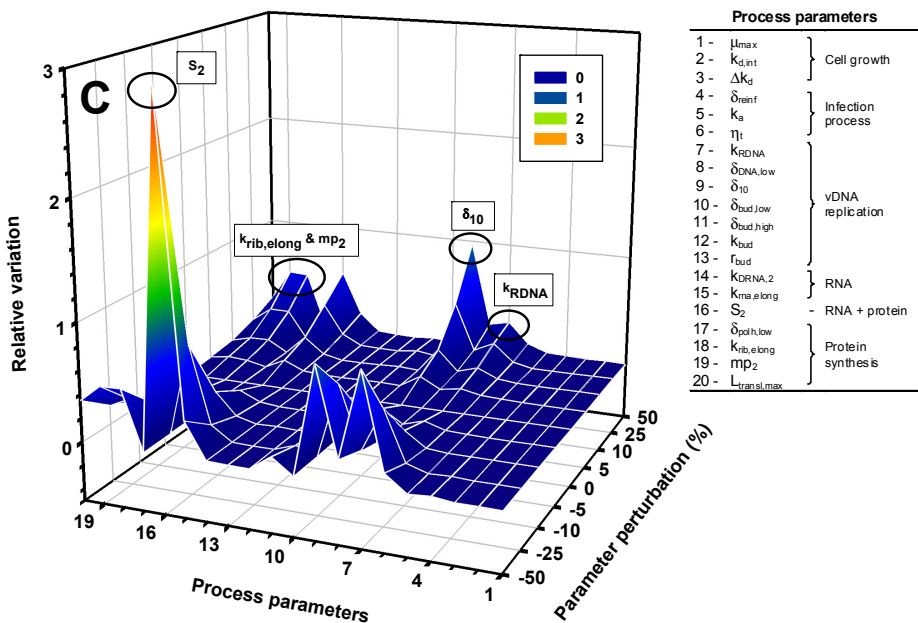
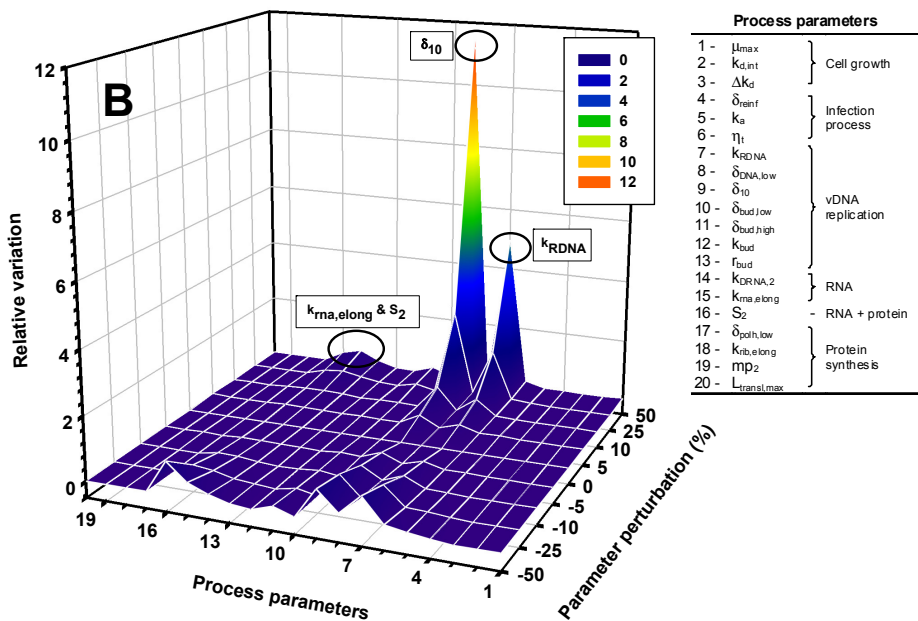


Interestingly, positive and negative perturbations in  $k_{RDNA}$  impact differently on the synthesis of progeny viruses. Virus budding rate seems to saturate at relatively low vDNA levels ( $1.5 \text{ dna.cell}^{-1}.\text{h}^{-1}$ ) suggesting that the bottleneck for virus budding is either the encapsidation machinery or cells are unable to secrete more than a specific number of viruses without affecting the integrity of the cellular membrane. Although the intracellular vDNA levels had varied with the magnitude of perturbation (Figure 36A), the saturation threshold was reached in both situations (positive and negative  $k_{RDNA}$  perturbations). Thus, the virus budding rate *per se* does not justify the aforementioned differences. Other factor must be also contributing, namely the time at which virus budding rate reached its saturation value. Indeed, this time was delayed when  $k_{RDNA}$  decreased as result of lower intracellular vDNA content. In contrast, no significant variation was observed when  $k_{RDNA}$  increased. Thus, the cumulative total of  $v_2$  ( $\text{virus.cell}^{-1}$  and  $\text{virus.ml}^{-1}$ ) is expected to vary significantly with the decreased in the value of  $k_{RDNA}$  but not with the increase in  $k_{RDNA}$ .

Figure 36A to 36C indicates that intracellular concentrations of vDNA, mRNA and viral protein are extremely sensitive to  $\delta_{10}$  and  $k_{RDNA}$ . The  $\mathbf{dna}_2$  variation with  $k_{RDNA}$  is linked to the vDNA replication pattern. Since vDNA replication follows first-order kinetics (consistent with template-limited replication), minor changes in the rate of vDNA synthesis propagates exponentially to the concentration of  $\mathbf{dna}_2$ . The  $\delta_{10}$  parameter influences vDNA concentration by scheduling the specific times for vDNA replication to start and stop. Knowing that vDNA replication follows first-order kinetics, perturbations in this time-window impacts deeply on  $\mathbf{dna}_2$ . Interestingly, the percentage of  $\mathbf{dna}_2$  variation seems to be correlated with the percentage of  $\delta_{DNA,low}$  variation; the higher the parameter perturbation, the higher is the output variation.

**Figure 36.** Sensitivity analysis of intracellular vDNA,  $\mathbf{dna}_2$  (A,  $\text{dna} \cdot \text{cell}^{-1}$ ), RNA,  $\mathbf{rna}_2$  (B,  $\text{rna} \cdot \text{ml}^{-1}$ ) and protein concentration,  $\mathbf{vp}_2$  (C,  $\mu\text{g} \cdot \text{cell}^{-1}$ ) with respect to model parameter perturbations (%) (0-12 - relative variation of model outputs upon parameter perturbation with respect to model outputs prior to parameter perturbation).

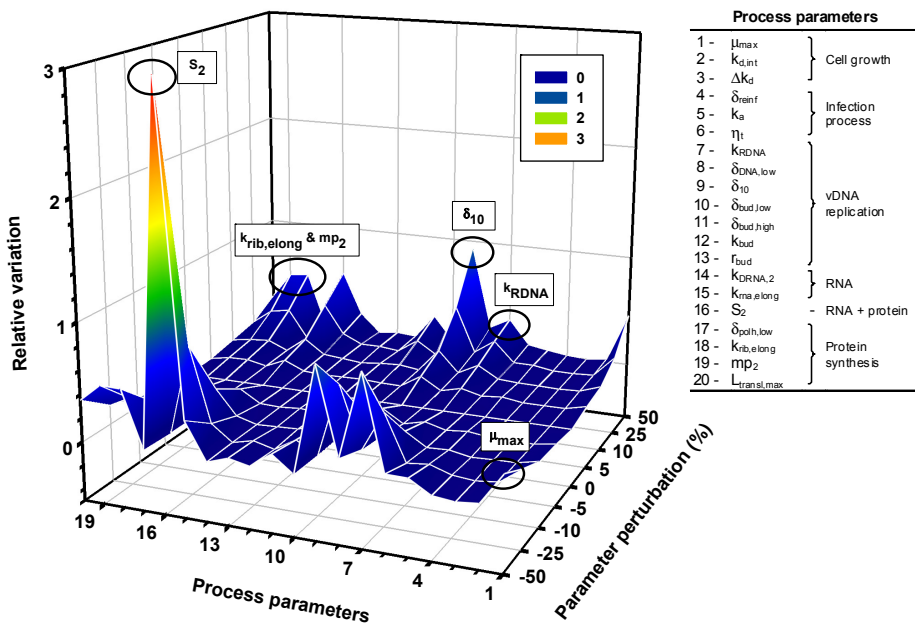




The dynamics of intracellular **rna<sub>2</sub>** and **vp<sub>2</sub>** follows closely **dna<sub>2</sub>** dynamics [13]; for this reason,  $\delta_{10}$  and  $k_{RDNA}$  have also a deep impact on **rna<sub>2</sub>** and **vp<sub>2</sub>**. Noteworthy is the higher influence of  $\delta_{10}$  and  $k_{RDNA}$  on final **rna<sub>2</sub>** levels

when compared to the influence of mRNA related parameters such as the gene size,  $S_2$ , and the RNA elongation rate,  $k_{rna,elong}$ . This suggests that  $rna_2$  concentration is more dependent on the replication levels of vDNA than on the efficiency of the transcriptional machinery.

The parameters with major influence on intracellular and total protein concentration are the gene size,  $S_2$ , the molecular weight of the viral protein,  $mp_2$ , and the ribosomal elongation rate,  $k_{rib,elong}$ , besides the abovementioned parameters  $\delta_{10}$  and  $k_{RDNA}$  (Figure 36C and 37). While almost irrelevant for intracellular protein expression, the  $\mu_{max}$  parameter has a noteworthy impact on total protein concentration.



**Figure 37.** Sensitivity analysis of total protein concentration,  $vp_2$  ( $\mu g \cdot ml^{-1}$ ) with respect to model parameter perturbations (%) (0-3 - relative variation of model outputs upon parameter perturbation with respect to model outputs prior to parameter perturbation).

Figures 34 to 37 are just examples of the amount of information that is possible to obtain about a given system when models such as the ones

presented in Chapters 3, 4 and 5 are used. The data generated is of paramount importance for the identification of system's critical parameters.

The sensitivity analysis here performed demonstrates that model outputs are more sensitive to changes in parameters related with vDNA and protein synthesis, such as  $\delta_{10}$ ,  $k_{RDNA}$ ,  $S_2$ ,  $mp_2$  and  $k_{rib,elong}$ , than to those linked to cell growth, infection process, virus budding, re-infection and mRNA.

It should be highlighted that these parameters have been estimated in Chapter 3 with high confidence intervals (see Table 18). Nevertheless, unaccounted regulatory mechanisms may impact in these parameter values at very different experimental conditions. If regulatory mechanisms are to be included in the future, they should focus on the list of most sensitive parameters above. An oriented design of experiments covering different regions of the process could be performed to tackle this issue.

## 4. OPTIMIZATION OF RLPS PRODUCTION

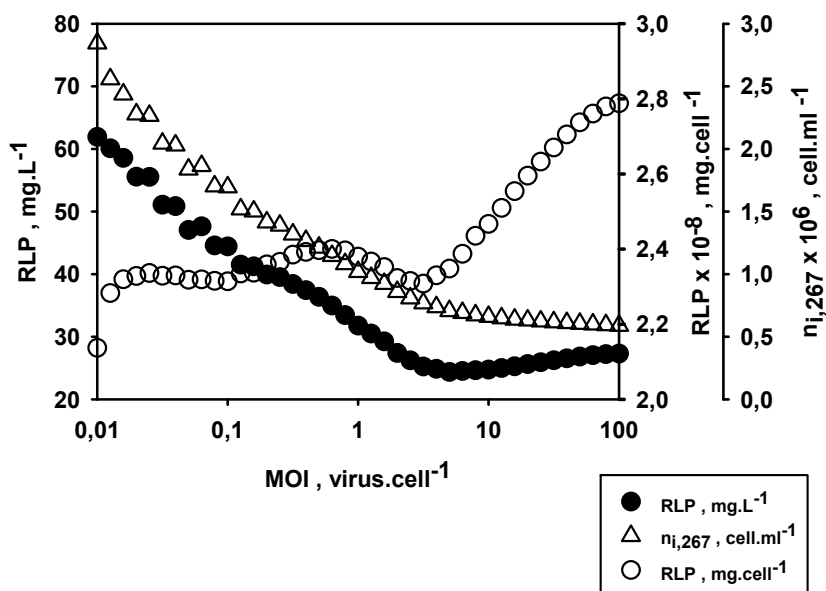
Optimal process and vector design parameters that maximise RLP productivity and assembly efficiency were investigated through model simulations. The optimisation degrees of freedom were the MOI and the promoter strength. Both single-infection and co-infection strategies were evaluated regarding their maximum productivity and assembly efficiency potential. Simulation time was 100 hpi for all scenarios presented below. Process performance was evaluated at time  $t=100$  hpi (end of process simulations).

### 4.1. Single-infection strategies

In single-infection strategies, the concentration of infected cells carrying the



genes coding for **vp2**, **vp6** and **vp7**,  $n_{i,267}$ , is identical to the overall concentration of infected cells,  $n_i$ . The effect of MOI on  $n_{i,267} = n_i$  is shown in Figure 38.



**Figure 38.** The effect of MOI in single-infection strategies: evaluating the intracellular ( $\text{mg.cell}^{-1}$ ) and total ( $\text{mg.L}^{-1}$ ) RLP production and the concentration of infected cells carrying the genes coding for **vp2**, **vp6** and **vp7**,  $n_{i,267}$  ( $\text{cell.ml}^{-1}$ ).

For synchronous infection ( $\text{MOI} > 2 \text{ virus.cell}^{-1}$ ), 100% infection occurs within the first couple of hours, meaning that maximum infected cell density is likely to be coincident with the cell concentration at infection ( $1 \times 10^6 \text{ cell.ml}^{-1}$ ). As time proceeds, infected cells began to die more rapidly as result of viral replication; the higher the viral load, the higher is the cell death rate. Thus, as MOI increases from 2 to 100  $\text{virus.cell}^{-1}$ ,  $n_{i,267}$  decreases accordingly (Figure 38). For asynchronous infection ( $\text{MOI} < 1 \text{ virus.cell}^{-1}$ ), only a small fraction of cells is infected initially; the lower the MOI, the lower is the concentration of infected cells. Uninfected cells will continue growing till the moment they are infected by newly formed viruses

that, in the meantime, bud from infected cells. Thus,  $n_{i,267}$  increases with the corresponding decrease in MOI.

The intracellular kinetics of RLP production converge to a maximum value, consistent with saturation in the translation machinery as described in Chapter 3 for **vp2** production (data not shown). This maximum was found to be a function of the MOI; the higher the MOI, the higher is the intracellular production of RLPs. The total concentration of RLPs depends more on the concentration of infected cells rather than on the intracellular production of RLPs. This explains why high intracellular RLP concentrations were not translated into high total RLP concentrations. Thus, the lower the MOI, the higher is  $n_{i,267}$  and the higher is the total concentration of RLPs.

Noteworthy is the independency of the assembly efficiency on the MOI (data not shown). This result is however rational because the genes coding for **vp2**, **vp6** and **vp7** are enclosed within the same baculovirus genome. Thus, the MOI is expected to have no influence on the ratios **vp6/vp2** and **vp7/vp2**; it only impacts on the expression levels of each viral protein. The estimated assembly efficiency was around 72%.

In conclusion, it does not bring any advantage to increase the MOI to very high levels in single-infection experiments due to the saturation of protein translation machinery, which has a lower impact on total RLP concentration than that of the concentration of infected cells which is maximised at low MOIs.

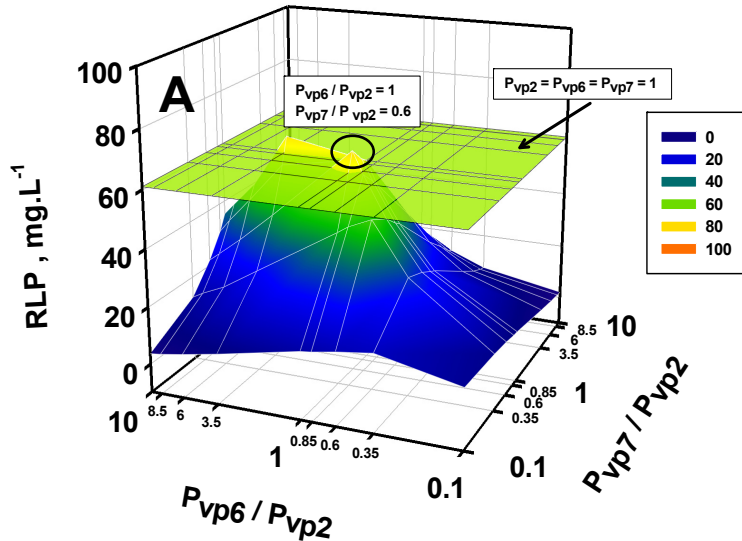
#### **4.1.1. The effect of promoter strength**

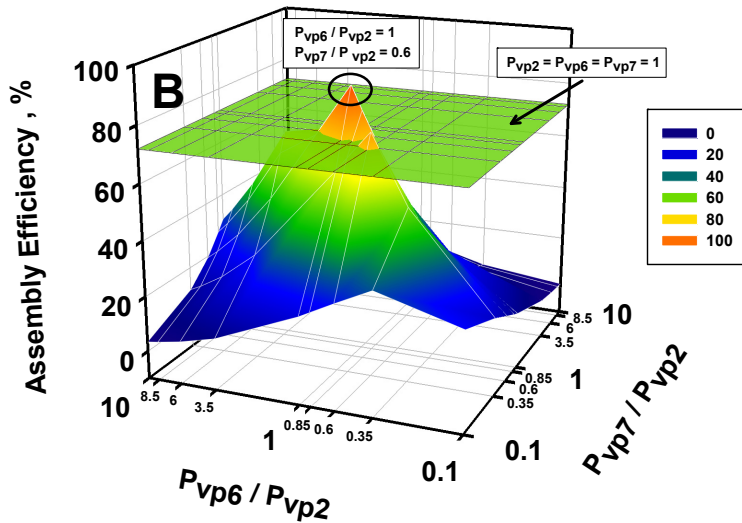
Optimisation of RLP production and assembly efficiency via vector design was investigated. More specifically, the effect of promoter strength on RLP assembly was studied at aid of model simulations. The strength of the

202

promoter controlling **vp2** expression ( $P_{vp2}$ ) was kept constant in all simulations whereas the promoter strength ratios  $P_{vp6}/P_{vp2}$  and  $P_{vp7}/P_{vp2}$  varied between 0.1 and 10. The transcription rate defined in Table 18 will be used to reflect these variations. For simplification, the case where all genes are under the control of promoters with equal strengths -  $P_{vp6}/P_{vp2} = P_{vp7}/P_{vp2} = 1$  - will be called from here on as “standard case”.

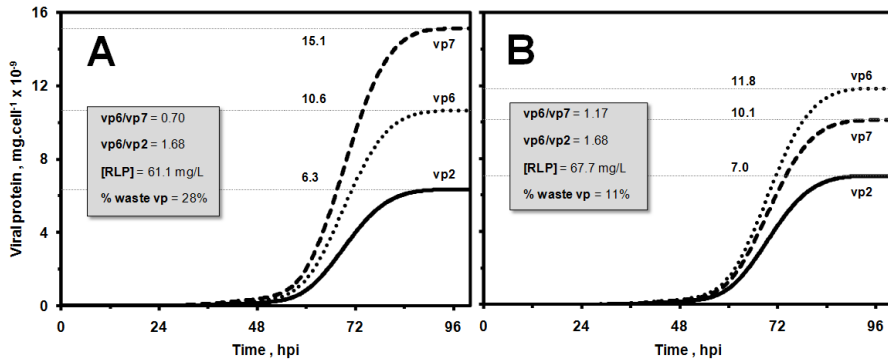
Variations in promoters' strength had a meaningless effect on the final concentration of infected cells (data not shown) because the strength of a promoter does not impact on cellular growth; it only affects the transcription and translation mechanisms. In contrast, they had a major impact on intracellular and total RLP production and assembly efficiency; a representative example is shown in Figure 39.





**Figure 39.** The effect of promoter strength on total RLP concentration (**A**,  $\text{mg.L}^{-1}$ ) and assembly efficiency (**B**, %) for single-infection strategies at  $\text{MOI}=0.01 \text{ virus.cell}^{-1}$ . The horizontal green surfaces are purely illustrative; they reflect the total RLP concentration ( $61 \text{ mg.L}^{-1}$ ) and the assembly efficiency (72%) achieved in the case where all genes are under the control of promoters with equal strengths -  $P_{vp2} = P_{vp6} = P_{vp7} = 1$ . They serve the purpose of aiding the reader in visualizing what conditions potentiate the increase in model outputs (region above the surface); below the surface, promoter strength variations have a negative impact on model output.

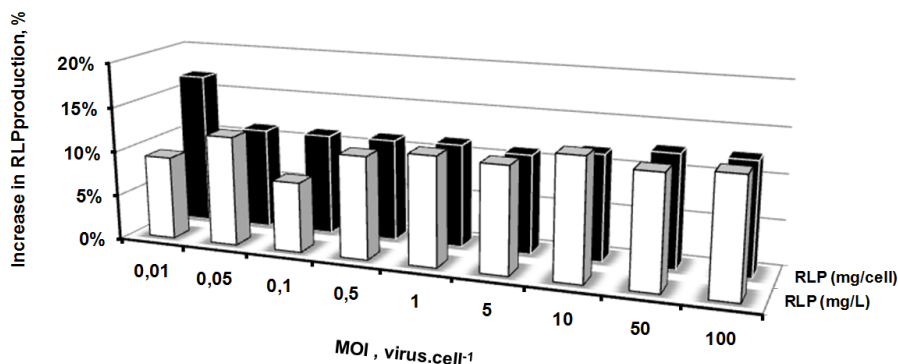
The combination of promoters' strength maximizing RLP production and assembly efficiency was  $P_{vp6}/P_{vp2} = 1$  and  $P_{vp7}/P_{vp2} = 0.6$ , irrespective of the MOI (see example in Figure 39). By decreasing  $P_{vp7}/P_{vp2}$  from 1 (standard case) to 0.6, the expression of **vp7** decreased in consequence of lower **rna<sub>7</sub>** concentration. As a result, the ratio **vp6/vp7** increased to levels close to the RLP stoichiometric ratio (1.21), meaning that few **vp7** and **vp6** were not assembled into RLPs. Although the ratio **vp6/vp2** had remained below the stoichiometric ratio of 2.84 (**vp6** is the limiting protein), the aforementioned combination of promoters' strength was the one minimizing more the total amount of unassembled proteins (see example in Figure 40).



**Figure 40.** The effect of promoter strength on intracellular protein expression ( $\text{mg}\cdot\text{cell}^{-1}$ ) for single-infection strategies at  $\text{MOI}=0.01$  virus. $\text{cell}^{-1}$ . Graph **A** represents the standard case ( $P_{vp2} = P_{vp6} = P_{vp7} = 1$ ) while graph **B** represents the infection strategy where  $P_{vp6}/P_{vp2} = 1$  and  $P_{vp7}/P_{vp2} = 0.6$ .

If **vp6** is the limiting protein in the standard case, one could speculate that  $P_{vp6}/P_{vp2}$  higher than 1 would even the relation between **vp6** and the other two proteins. However, since maximum translation capacity is a function of RNA levels (see Table 18), the increase in  $P_{vp6}$  induces higher translation machinery utilization, positively impacting on the concentration of **vp6** but strongly inhibiting the expression of **vp2** and **vp6**. Consequently, RLP production and assembly efficiency are compromised. These results suggest that the balance between promoter strength and translation machinery utilization is critical for maximum process performance.

Figure 41 depicts the maximum increase in intracellular and total RLP production that can be achieved through genetic redesign for a selective range of MOIs studied in Figure 38. In other words, for each MOI, it compares the best case achieved by manipulation of promoters' strength with the standard case.

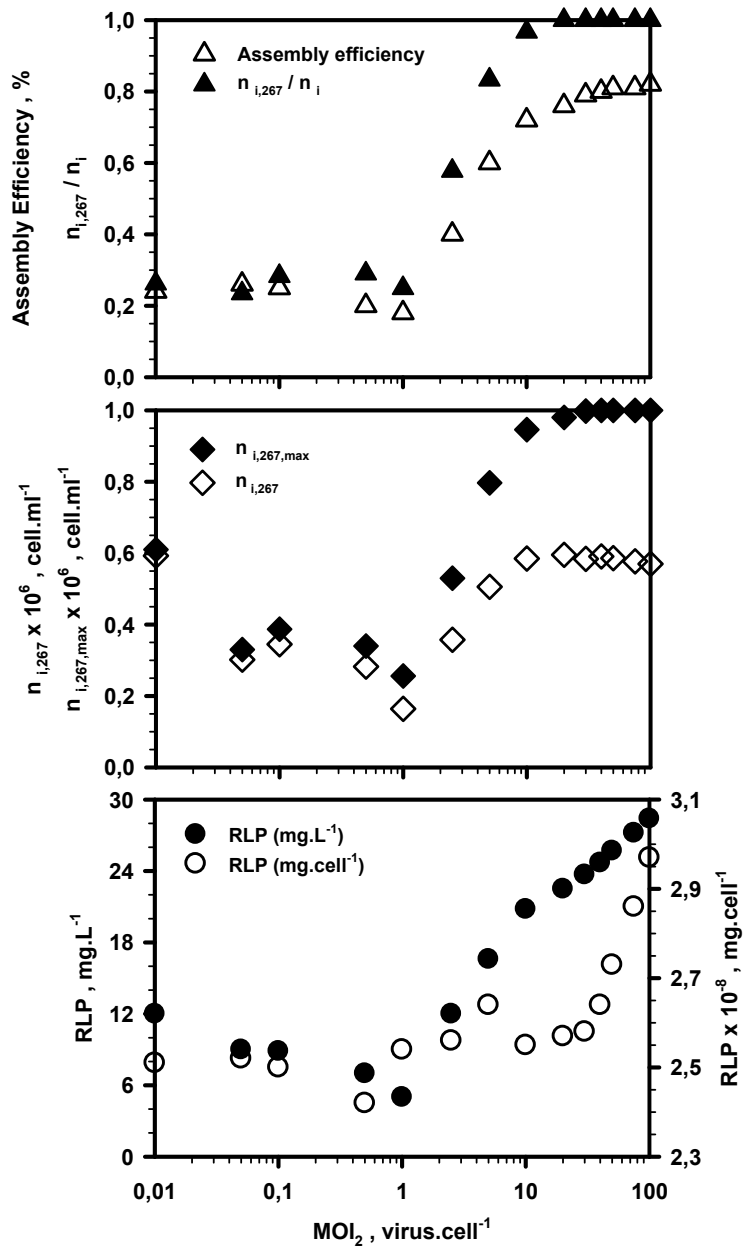


**Figure 41.** The effect of promoter strength on intracellular ( $\text{mg}\cdot\text{cell}^{-1}$ ) and total ( $\text{mg}\cdot\text{ml}^{-1}$ ) RLP production for single-infection strategies: maximum increase in RLP production (%) that can be achieved through genetic redesign for a selective range of MOIs.

An improvement between 7% and 18% in RLP production and around 24% in assembly efficiency was achieved when compared to the standard case. Although configuring a somewhat modest improvement potential, these data demonstrate that the redesign of baculovirus vectors, in which the promoters controlling viral protein expression can be selected according to their strength, can improve RLP production and minimize the amount of unassembled proteins.

#### 4.2. Co-infection strategies

The effect of MOI on cell growth, RLP production and assembly efficiency in co-infection strategies was investigated through model simulations. The number of monocistronic baculoviruses coding for **vp2** per cell (**MOI<sub>2</sub>**) used for process simulations varied from 0.01 to 100  $\text{virus}\cdot\text{cell}^{-1}$ ; the number of monocistronic baculoviruses coding for **vp6** and **vp7** per cell, **MOI<sub>6</sub>** and **MOI<sub>7</sub>** respectively, were defined so that **MOI<sub>6</sub>/MOI<sub>2</sub>** and **MOI<sub>7</sub>/MOI<sub>2</sub>** could vary between 0.01 and 10. For each **MOI<sub>2</sub>**, the optimal **MOI<sub>6</sub>** and **MOI<sub>7</sub>** maximizing process outputs were selected for representation in Figure 42.



**Figure 42.** The effect of MOI on intracellular ( $\text{mg.cell}^{-1}$ ) and total ( $\text{mg.ml}^{-1}$ ) RLP production, assembly efficiency (%), concentration of infected cells carrying the genes coding for **vp2**, **vp6** and **vp7**,  $n_{i,267}$  ( $\text{cell.ml}^{-1}$ ) and maximum  $n_{i,267}$ ,  $n_{i,267,max}$  ( $\text{cell.ml}^{-1}$ ) for co-infection strategies. The overall infected cell concentration is denoted by  $n_i$ . For each  $MOI_2$ , the optimal  $MOI_6$  and  $MOI_7$  maximizing process outputs were selected for representation.

As expected, the probability of a cell being infected by the three monocistronic baculoviruses was found to be highly dependent on the MOI (Figure 42). The sigmoidal curve of  $n_{i,267}/n_i$  as function of MOI shows that high MOIs favours the synchronous infection of insect cells with the three monocistronic baculoviruses, mostly as a consequence of the high initial concentration of baculoviruses; thus,  $n_{i,267}/n_i$  converged to 1 and maximum  $n_{i,267}$  values ( $n_{i,267,max}$ ) coincided with the initial uninfected cell concentration ( $1 \times 10^6$  cells.ml<sup>-1</sup>). As the MOI decreases ( $< 2$  virus.cell<sup>-1</sup>), the probability of a cell population not being infected with the three recombinant baculoviruses increases due to the random nature of virus binding to the host cells surface. For these reason, the ratio  $n_{i,267}/n_i$  as well as the  $n_{i,267,max}$  are significantly lower than the ones observed for high MOIs. The apparent local minimum of  $n_{i,267,max}$  at  $MOI_2=1$  virus.cell<sup>-1</sup> can be explained by the following factors: the MOI is insufficient to ensure that all cells are infected by the three baculoviruses but sufficiently high for 100% infection in the first infection cycle. By the time that virus budding commences, re-infection is no longer possible; hence, most of the cells will remain infected by only one or two baculoviruses, which significantly reduces  $n_{i,267,max}$ .

The profiles of  $n_{i,267}$  and  $n_{i,267,max}$  with respect to MOI are very similar. At low MOIs they almost coincide, meaning that maximum concentration of cells carrying the genes coding for the three viral proteins are achieved at the end of process simulations. At high MOIs ( $> 2$  virus.cell<sup>-1</sup>) they differ significantly; the lower  $n_{i,267}$  values are related with the “metabolic burden” effect caused by the high intracellular viral load. Therefore, the advantage of using low MOIs for achieving high infected cell concentrations, as in single infection experiments, turns out to be useless in co-infection experiments as the probability of a host cell carrying the three genes coding for **vp2**, **vp6** and **vp7**, proteins necessary to build the RLP, is very low.



The intracellular RLP concentration profile follows closely the one of single-infection (compare Figure 38 and Figure 42); upon an apparent plateau for  $\text{MOI} < 10 \text{ virus.cell}^{-1}$ , intracellular RLP production increases with the increase in the MOI. The profile of total RLP concentration is different. At  $\text{MOIs} < 10 \text{ virus.cell}^{-1}$ , the distinguish factor in RLP production is  $n_{i,267}$ ; thus, the profile of total RLP concentration resembles the one observed for  $n_{i,267}$  and  $n_{i,267,\text{max}}$ . At higher MOIs, the distinguish factor is the intracellular RLP production; hence, the profile follows closely the one of intracellular RLP production. Most importantly, Figure 42 indicates that, as opposed to single infection experiments, the total concentration of RLPs increases with the increase in MOI in co-infection experiments.

The assembly efficiency was found to be correlated with the MOI: it increases with increasing MOI due to the aforementioned random infection effect (see  $n_{i,267}/n_i$  correlation with MOI) with a local minimum at  $\text{MOI}_2=1$  (Figure 42). At  $\text{MOI}_2$ ,  $\text{MOI}_6$  and  $\text{MOI}_7 < 2 \text{ virus.cell}^{-1}$ , the percentage of cells expressing either one or two viral proteins is around 75%, contrasting with the residual value observed at higher MOIs. Thus, the amount of unassembled proteins is expected to be much higher, drastically reducing the assembly efficiency. The local minimum at  $\text{MOI}_2=1 \text{ virus.cell}^{-1}$  is directly linked to the minimum observed for  $n_{i,267,\text{max}}$ .

For the  $\text{MOI}_2$  evaluated, the combination of  $\text{MOI}_6$  and  $\text{MOI}_7$  maximizing RLP production and assembly efficiency varied considerably. For  $\text{MOI}_2$ ,  $\text{MOI}_6$  and  $\text{MOI}_7$  higher than  $2 \text{ virus.cell}^{-1}$ , these variables are maximized whenever:

- $\text{MOI}_6/\text{MOI}_2 \approx 2.1$
- $\text{MOI}_7/\text{MOI}_2 \approx 1.2$

At these conditions, not only all cells are infected synchronously and by at

least one recombinant baculovirus coding for **vp2**, **vp6** and **vp7** ( $n_{i,267} \rightarrow n_i$ ) but also the expression levels of the three viral proteins almost obey to the stoichiometry of the particle, thus reducing the amount of waste proteins. At low MOIs (**MOI<sub>2</sub>**, **MOI<sub>6</sub>** and **MOI<sub>7</sub>** < 2 virus.cell<sup>-1</sup>), there is no clear trend in the **MOI<sub>6</sub>/MOI<sub>2</sub>** and **MOI<sub>7</sub>/MOI<sub>2</sub>** ratios, which is intrinsically related with the high degree of stochasticity of the infection process at such low MOIs.

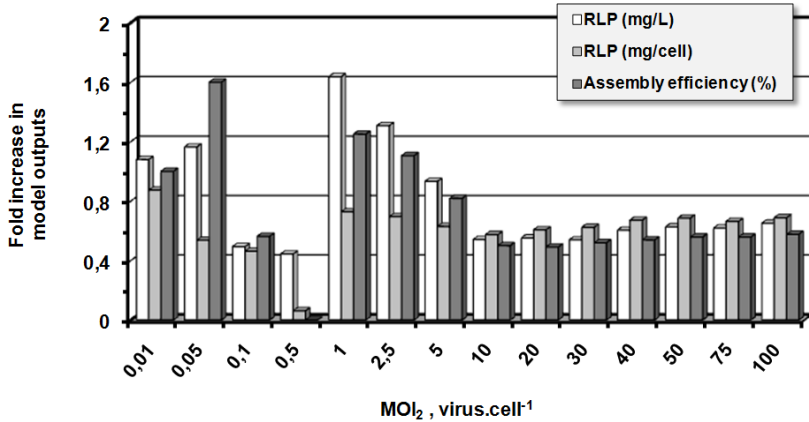
In conclusion, these results demonstrate that high MOIs maximize RLP production at co-infection strategies, which is precisely the opposite trend verified for the single infection strategy.

#### 4.2.1. Are traditional co-infection strategies suboptimal?

The production of RLPs is commonly performed using identical number of recombinant baculoviruses coding for **vp2**, **vp6** and **vp7** (**MOI<sub>2</sub>** = **MOI<sub>6</sub>** = **MOI<sub>7</sub>**). This strategy was compared to optimal MOI combinations of individual baculoviruses in terms of relative RLP production and assembly efficiency.

It is evident from Figure 43 that traditional strategies are far-off from being the ideal ones; intracellular and total RLP concentrations and assembly efficiency are significantly lower, independently of the MOI. In fact, at the MOIs normally used for RLP production (**MOI<sub>2</sub>** = **MOI<sub>6</sub>** = **MOI<sub>7</sub>** = 3 to 5 virus.cell<sup>-1</sup>), model simulations show that a 1-fold increase in total RLP concentration, the most important performance criterion, is achievable by oriented-manipulation of the **MOI<sub>2</sub>**, **MOI<sub>6</sub>** and **MOI<sub>7</sub>**.

These results highlight the window of productivity gains that are possible through model-based optimisation for design of novel optimal infection strategies.



**Figure 43.** Comparison of model outputs for traditional co-infection strategies ( $\mathbf{MOI_2=MOI_6=MOI_7}$ ) and model-designed strategies based on oriented-manipulation of the MOI of individual baculoviruses. The fold increase denotes the increase in intracellular ( $\text{mg}\cdot\text{cell}^{-1}$ ) and total ( $\text{mg}\cdot\text{ml}^{-1}$ ) RLP production and in assembly efficiency (%) by oriented-manipulation of the MOI of individual baculoviruses in relation to traditional co-infection strategies.

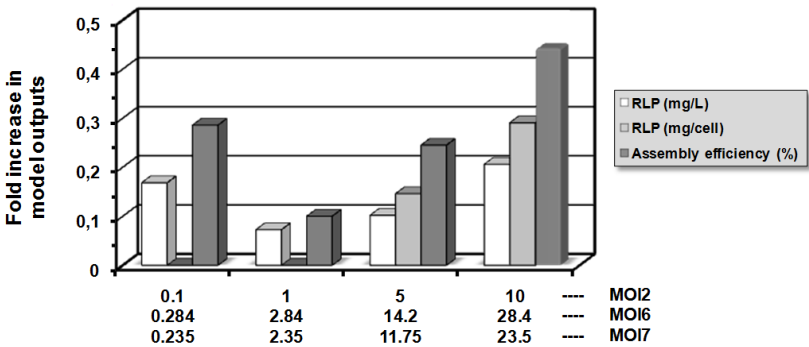
#### 4.2.2. Promoter effect at co-infection strategies

To test whether, at co-infection strategies, promoter strength impacts or not on RLP production, different scenarios were evaluated: the strength of the promoter controlling **vp2** expression ( $\mathbf{P_{vp2}}$ ) was maintained constant in all simulations whereas  $\mathbf{P_{vp6}/P_{vp2}}$  and  $\mathbf{P_{vp7}/P_{vp2}}$  were varied between 0.1 and 10. For a selected combination of  $\mathbf{MOI_2}$ ,  $\mathbf{MOI_6}$  and  $\mathbf{MOI_7}$ , the optimal combination of promoters' strengths in terms of relative RLP production and assembly efficiency was compared to the co-infection strategy where all genes are under the control of promoters with equal strengths ( $\mathbf{P_{vp6}/P_{vp2} = P_{vp7}/P_{vp2} = 1}$ ) and the improvement in model outputs calculated (Figure 44).

No significant variation in  $\mathbf{n_{i,267}}$  and  $\mathbf{n_{i,267}/n_i}$  was observed for the  $\mathbf{MOI_2 = 5}$  and 10  $\text{virus}\cdot\text{cell}^{-1}$ ; at lower MOIs, there was an increase in  $\mathbf{n_{i,267}}$  and  $\mathbf{n_{i,267}/n_i}$  which is linked to the stochasticity of the infection process (data not shown). This is rational because the promoter strength does not impact

directly or indirectly (via vDNA replication rate or synthesis and budding of progeny viruses) on cellular growth; it only affects the transcription and translation mechanisms.

Optimal combination of promoters' strengths had a positive impact on intracellular RLP production at  $\text{MOI}_2 = 5$  and  $10 \text{ virus.cell}^{-1}$  (Figure 44). Indeed, by controlling the relative strength of each promoter ( $P_{vp6}/P_{vp2}$  and  $P_{vp7}/P_{vp2}$ ), there was a better adjustment of the expression levels of the three viral proteins to what thermodynamically favours RLP assembly. This reduced the amount of unassembled proteins and increased not only the assembly efficiency but also the total RLP concentration. At  $\text{MOI}_2 = 0.1$  and  $1 \text{ virus.cell}^{-1}$ , no significant variation in intracellular RLP concentration was observed; hence, the increase in total RLP concentration and assembly efficiency is the result of the increase in  $n_{i,267}$ .



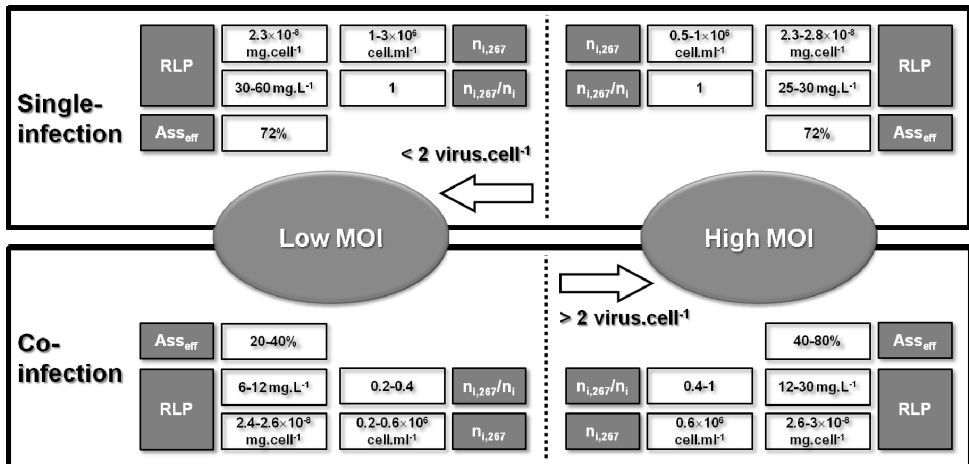
**Figure 44.** Comparison of model outputs (intracellular ( $\text{mg.cell}^{-1}$ ) and total ( $\text{mg.ml}^{-1}$ ) RLP production and assembly efficiency (%)) for co-infection strategies where the genes coding for the three viral proteins are under the control of promoters with equal strengths ( $P_{vp6}/P_{vp2} = P_{vp7}/P_{vp2} = 1$ ) and co-infection strategies maximizing RLP production based on the manipulation of the strength of each promoter. The fold increase denotes the increase in model outputs by oriented-manipulation of the strength of each promoter.

Summing up, the redesign of baculovirus vectors in co-infection strategies, in which the promoters controlling viral protein expression can be selected

according to their strength, is a potential route for optimizing RLP production and assembly efficiency.

#### 4.3. What to choose: single-infection or co-infection?

Model simulations indicate that single-infections give better results than co-infections. This is clearly evident for low MOIs (Figure 45). Since single-infections are more competent than co-infections in guaranteeing that infected cells carry all genes required for the expression of the three viral proteins, significant differences in  $n_{i,267}$ ,  $n_{i,267}/n_i$  and total RLP concentration were observed. Intracellular RLP production and assembly efficiency were not so strongly influenced by the stochasticity of the infection process; nonetheless, variations as high as 67% were observed. At high MOIs, the difference between single- and co-infection outputs is meaningless. This shows that the statistics of infection is no longer a key factor; RLP production becomes controlled by the expression levels of each viral protein.



**Figure 45.** Single-infection vs co-infection: the effect of MOI on  $n_{i,267}$  (concentration of infected cells carrying the genes coding for **vp2**, **vp6** and **vp7**),  $n_{i,267}/n_i$  (ratio between  $n_{i,267}$  and the overall infected cell concentration,  $n_i$ ), RLP production (**RLP**) and assembly efficiency (**Ass<sub>eff</sub>**).

By redesigning baculovirus vectors in single-infection strategies, in which the promoters controlling viral protein expression can be selected according to their strength, it is possible to improve RLP production even further (see section 4.1.1). In conclusion, these findings clearly demonstrate that single-infections offer a greater potential for RLP process optimization than co-infection strategies.

## 5. CONCLUSIONS AND FUTURE PERSPECTIVES

From the model sensitivity analysis and RLP production optimisation study performed in this chapter, the following main conclusions can be outlined:

### Sensitivity analysis

- vDNA and protein related parameters, such as  $\delta_{10}$ ,  $k_{RDNA}$ ,  $S_2$ ,  $mp_2$  and  $k_{rib,elong}$ , are much more relevant than cell growth and mRNA parameters for protein expression - model outputs are highly sensitive to changes in these parameters;
- Variations in parameters related with the infection process, virus budding and re-infection have a minor influence on model outputs;

### RLP optimization: single-infection

- The concentration of infected cells carrying the genes coding for **vp2**, **vp6** and **vp7**,  $n_{i,267}$ , matches the overall concentration of infected cells,  $n_i$ , irrespective of the MOI;
- $n_{i,267}$  is higher at low MOIs ( $1-3 \times 10^6 \text{ cell.ml}^{-1}$ ) than at high MOIs ( $0.5-1 \times 10^6 \text{ cell.ml}^{-1}$ );
- The specific concentration of RLPs is function of the MOI – it

increases from  $2.3 \times 10^{-8}$  mg.cell<sup>-1</sup> (low MOIs) to  $2.8 \times 10^{-8}$  mg.cell<sup>-1</sup> (high MOIs);

- The total concentration of RLPs decreases from 60 mg.L<sup>-1</sup> to 25 mg.L<sup>-1</sup> with the increase in the MOI;
- The assembly efficiency is independent of the MOI, being around 72%;
- The strength of the promoters controlling the expression of **vp2**, **vp6** and **vp7** have no influence on  $n_{i,267}$ ; in contrast, optimal combination of promoters' strengths induces an improvement in intracellular and total RLP production (7% to 18%) and in assembly efficiency (24%);
- In summary, process optimization in single-infection strategies can be achieved by manipulation of the MOI (lower MOIs) and/or by redesigning baculovirus vectors, in which the promoters controlling viral protein expression can be selected according to their strength (lowering the strength of the promoter controlling **vp7** expression in relation to the other two promoters);

### RLP optimization: co-infection

- High MOIs favours the infection of insect cells with the three monocistronic baculoviruses ( $n_{i,267}/n_i \rightarrow 1$ ); as the MOI decreases, the infection process becomes stochastic and cells are likely to be infected by only one or two of the three recombinant baculoviruses (minimum of  $n_{i,267}/n_i$  at 0.2);
- Intracellular RLP production follows a two-step profile: upon an apparent plateau for MOIs  $< 10$  virus.cell<sup>-1</sup> ( $2.4\text{-}2.6 \times 10^{-8}$  mg.cell<sup>-1</sup>),

intracellular RLP concentrations increase with the increase in the MOI ( $2.6\text{--}3\times 10^{-8}$  mg.cell<sup>-1</sup>);

- Total RLP production increases with the MOI (6-30 mg.L<sup>-1</sup>);
- The assembly efficiency is correlated with the MOI: there is a local minimum at **MOI<sub>2</sub>** = 1 virus.cell<sup>-1</sup> (20%); for lower MOIs, the efficiencies vary between 20% and 40%; for higher MOIs, the assembly efficiency increases to 80%;
- For **MOI<sub>2</sub>**, **MOI<sub>6</sub>** and **MOI<sub>7</sub>** higher than 2 virus.cell<sup>-1</sup>, RLP production and assembly efficiency are maximized whenever **MOI<sub>6</sub>/MOI<sub>2</sub>**  $\approx$  2.1 and **MOI<sub>7</sub>/MOI<sub>2</sub>**  $\approx$  1.2; at lower MOIs, there is no clear trend in the **MOI<sub>6</sub>/MOI<sub>2</sub>** and **MOI<sub>7</sub>/MOI<sub>2</sub>** ratios maximizing process outputs;
- In summary, high MOIs maximize RLP production; the redesign of baculovirus vectors configures a somewhat modest improvement potential;

### **RLP optimization: single-infection or co-infection?**

At low MOIs, single-infection strategies offer significant advantages over co-infection experiments, namely:

- **n<sub>i,267</sub>** can reach up to  $3\times 10^6$  cell.ml<sup>-1</sup>, contrasting with the  $0.6\times 10^6$  cell.ml<sup>-1</sup> at co-infection;
- Higher total RLP concentrations: 30-60 mg.L<sup>-1</sup> against the 6-12 mg.L<sup>-1</sup> at co-infection;
- Assembly efficiencies are higher: 72% against the 20-40% at co-infection;



At high MOIs, the differences in model outputs for co- and single-infection strategies are meaningless:

- $n_{i,267}$  levels are within  $0.5\text{-}1 \times 10^6 \text{ cell.ml}^{-1}$ ;
- Maximum total RLP concentration is around  $30 \text{ mg.L}^{-1}$ ;
- Maximum assembly efficiencies are between 72-80%.

The findings described throughout Chapter 6 demonstrate that the redesign of baculovirus vectors, in which the promoters controlling viral protein expression can be selected according to their strength, and the manipulation of the MOI have great potential for process optimization. The practical implementation of these concepts may not be easy since it depends on the available resources, namely the type of recombinant baculoviruses (tracistronic or monocistronic), and on the knowledge of molecular biology techniques, essential for the construction of novel recombinant baculoviruses.

## 6. ACKNOWLEDGEMENTS

This work was supported by European Commission project (QLRT-2001-01249) and the Portuguese Fundação para a Ciência e Tecnologia (POCTI/BIO/55975/2004 and SFRH/BD/21910/2005).

## 7. REFERENCES

1. Mattion, N.M., Cohen, J., and Estes, M.K., **1994**. The rotavirus proteins. Viral infections of the gastrointestinal tract, ed. Kapikian, A. Marcel Dekker, Inc. 169-249. New York, USA.
2. Crawford, S.E., *et al.*, **1994**. Characterization of virus-like particles produced by the expression of rotavirus capsid proteins in insect cells. *J Virol* 9 (68): 5945-52.

3. Sabara, M., *et al.*, **1991**. Assembly of double-shelled rotaviruslike particles by simultaneous expression of recombinant VP6 and VP7 proteins. *J Virol* 12 (65): 6994-7.
4. Crawford, S.E., *et al.*, **1999**. Heterotypic protection and induction of a broad heterotypic neutralization response by rotavirus-like particles. *J Virol* 6 (73): 4813-4822.
5. Cruz, P.E., *et al.*, **1998**. Optimization of the production of virus-like particles in insect cells. *Biotechnol Bioeng* 4 (60): 408-18.
6. Jiang, B., *et al.*, **1999**. Heterotypic protection from rotavirus infection in mice vaccinated with virus-like particles. *Vaccine* 7-8 (17): 1005-13.
7. Maranga, L., Brazao, T.F., and Carrondo, M.J., **2003**. Virus-like particle production at low multiplicities of infection with the baculovirus insect cell system. *Biotechnol Bioeng* 2 (84): 245-53.
8. Mellado, M.C.M., *et al.*, **2009**. Assessment of RLPs production and assembly efficiency using capillary zone electrophoresis. In manuscript.
9. Palomares, L.A., Lopez, S., and Ramirez, O.T., **2002**. Strategies for manipulating the relative concentration of recombinant rotavirus structural proteins during simultaneous production by insect cells. *Biotechnol Bioeng* 6 (78): 635-44.
10. Roldão, A., *et al.*, **2007**. Modeling rotavirus-like particles production in a baculovirus expression vector system: Infection kinetics, baculovirus DNA replication, mRNA synthesis and protein production. *J Biotechnol* 4 (128): 875-894.
11. Roldão, A., *et al.*, **2006**. Intracellular dynamics in rotavirus-like particles production: Evaluation of multigene and monocistronic infection strategies. *Process Biochemistry* 10 (41): 2188-2199.
12. Vieira, H., *et al.*, **2005**. Triple layered rotavirus VLP assembly: kinetics of vector replication, mRNA stability and recombinant protein production. *J Biotechnol*, 1 (120): 72-82.
13. Roldão, A., *et al.*, **2008**. Stochastic simulation of protein expression in the baculovirus/insect cells system. *Computers & Chemical Engineering* 1-2 (32): 68-77.
14. Wang, M.Y., Kwong, S., and Bentley, W.E., **1993**. Effects of oxygen/glucose/glutamine feeding on insect cell baculovirus protein expression: a study on epoxide hydrolase production. *Biotechnol Prog* 4 (9): 355-61.
15. Dalal, N.G. and Bentley, W.E., **1999**. Mathematical characterization of insect cell (Sf-9) death in suspended culture. *Biotechnol Lett* (21): 325-329.
16. Hu, Y.-C. and Bentley, W.E., **2000**. A kinetic and statistical-thermodynamic model for baculovirus infection and virus-like particle assembly in suspended insect cells. *Chem Eng Sci* (55): 3991-4008.
17. Dee, K.U. and Shuler, M.L., **1997**. Optimization of an assay for baculovirus titer and design of regimens for the synchronous infection of insect cells. *Biotechnol Prog* 1 (13): 14-24.
18. Dee, K.U. and Shuler, M.L., **1997**. A mathematical model of the trafficking of acid-dependent enveloped viruses: Application to the binding, uptake, and nuclear accumulation of baculovirus. *Biotechnol Bioeng* 5 (54): 468-490.
19. Blissard, G.W. and Rohmann, G.F., **1990**. Baculovirus diversity and molecular biology. *Annu Rev Entomol* (35): 127-55.
20. Tjia, S.T., zu Altschiltschesche, G.M., and Doerfler, W., **1983**. Autographa californica nuclear polyhedrosis virus (AcNPV) DNA does not persist in mass cultures of mammalian cells. *Virology* 1 (125): 107-17.
21. Hu, Y.C. and Bentley, W.E., **2001**. Effect of MOI ratio on the composition and yield of chimeric infectious bursal disease virus-like particles by baculovirus co-infection: deterministic predictions and experimental results. *Biotechniques* 1 (75): 104-119.
22. Possee, R.D. and King, L.A., **1992**. The Baculovirus Expression System: A laboratory guide, ed. Hall, C. London.
23. Labbe, M., *et al.*, **1991**. Expression of rotavirus VP2 produces empty corelike particles. *J Virol* 6 (65): 2946-52.

# Chapter 7

**D**iscussion and conclusions

**CONTENTS**

**1. Discussion .....221**

1.1. MOI controllability .....221

1.2. Model-based optimisation of the MOI for single protein expression .....222

1.3. Model-based optimisation of RLP production .....224

1.3.1. RLP assembly controllability .....224

1.3.2. RLP optimization by oriented-manipulation of the MOI .....226

1.3.3. RLP optimization via genetic redesign .....227

1.3.4. Single infection or co-infection: what to choose? .....229

**2. Conclusions.....230**

**3. Future work.....233**

**4. References .....235**

## 1. DISCUSSION

In previous Chapters, *in silico* baculovirus/insect cells optimisation tools for heterologous protein production and for virus like particles production were developed. The target product was rotavirus-like-particles (RLP), one of the most promising immunization strategies against rotavirus disease. RLP is representative of many other biotherapeutic candidates with major production hurdles. The extensive production of this vaccine candidate has been hindered essentially due to the complexity of the RLP production process, which commonly leads to very low yields and volumetric productivities. In this Chapter, it is highlighted the value added but also the main pitfalls of the computational tools developed in previous Chapters for the optimisation of the baculovirus/insect cells system.

### 1.1. MOI controllability

The ability to control the multiplicity of infection (MOI) is a mandatory guideline for ensuring batch consistency. Without accurate MOI control, optimisation of the baculovirus/insect cell system for heterologous protein expression becomes compromised. This implies accurate estimation of baculovirus titers so that the MOI measurement encloses as little variability as possible. In order to identify the titration methods that better combine accuracy with the method's cost and work intensity, a comparative study was performed in Chapter 2.

The endpoint dilution, microculture tetrazolium and flow cytometric assays presented the best compromise between cost, time and accuracy. Their cost was similar and, in average, 75% lower than those reported for the alamarblue, real time Q-PCR and plaque assays. This analysis was based on the cost of consumables; costs associated with equipment and human resources were not considered since these differ from lab to lab.

The standard errors were below the 28-36% reported for the alamarblue and growth cessation assays but slight above the ones achieved for the plaque assay and real time Q-PCR (7-10%). These values are in agreement with literature values and within acceptable variations defined by regulatory agencies [1-3]. The titration time is their main disadvantage; nonetheless, it may be viewed as a reasonable price to pay for estimating accurate titers.

Interestingly, real time Q-PCR was found to be one of the techniques estimating viral titers statistically different from the average. The presence of defective interfering particles (DIPs) in viral stocks is the main cause. Since it is impossible to discriminate between infectious particles and DIPs with this technique, both viruses were counted as infectious and, thus, titers were over-estimated. This demonstrates how important it is to control the passage number of baculoviruses. Indeed, although their effect becomes more pronounced from passage 20 onwards, DIPs can be detected as soon as passage 10 and reach up to 60% of total replicated viruses [4].

Thus it may be realistic to expect a MOI variability of around 20%.

## **1.2. Model-based optimisation of the MOI for single protein expression**

To investigate the implications of MOI on recombinant protein expression in the baculovirus/insect cell system, a stochastic structured model was developed in Chapter 3.

The proposed model represents an advance over previous modelling studies [5-7] since it combines an explicit stochastic infection process with mass action kinetics describing the dynamics of vDNA, mRNA and foreign protein synthesis. In addition, it explores the hypothesis of gene size being the main factor affecting the rate of protein expression in the host cell.

Model simulations confirmed that optimal MOIs are typically low, in agreement with experimental observations and other modelling studies [6; 8]. Indeed, at MOIs lower than 1 virus.cell<sup>-1</sup> protein expression was maximized as a result of higher infected cell concentration at the end of bioreaction - asynchronous infection [9]. On the other extreme, for MOIs higher than 1 virus.cell<sup>-1</sup>, lower protein concentrations were achieved in consequence of protein expression saturation and low infected cell concentration – synchronous infection [9]. This means that it does not bring any advantage to increase the MOI to very high levels as the translation machinery saturates, thus resulting in no improvement in protein expression.

Chapter 3 results clearly indicate that intracellular vDNA content reaches a maximum level which was found to be a trade-off between high MOIs (more ‘seeds’ for vDNA replication with reduced time window for its replication) and extended virus budding effect at low MOIs (more virus for infection over a longer period of time). The time at which it occurs is coincident with the halt in mRNA and protein synthesis, and inversely proportional to the MOI. The higher the MOI, the more severe is the “metabolic burden” and, consequently, the lower is the time window for intracellular vDNA, mRNA and protein synthesis, in agreement with the literature [6; 10-12].

An interesting result is that the gene size did not affect the optimal MOI for protein expression. This is reasonable because the recombinant gene size does not affect either the infection kinetics or the vDNA replication rate or the budding of progeny virus. It only affects the transcription and translation mechanisms, meaning that it only impacts on the intracellular accumulation of protein. This means that different proteins are expected to have optimal productivities around the same MOI obtained for vp2:

- MOI = 0.01 – 0.1 virus.cell<sup>-1</sup>

On the other hand, gene size impacts negatively on protein expression as showed elsewhere [6]; the larger the gene size, the lower is the final protein concentration.

A final point is the effect of MOI errors on protein optimisation. Taking the example of vp2 and a typical error in MOI of 20% around the optimum value of  $0.01 \text{ virus.cell}^{-1}$ , a degradation between 5% and 30% in the concentration of viral protein is expected (see Figure 19 in Chapter 3). These results emphasize the importance of accurate titration methods for the controllability of the MOI and, consequently, protein expression.

### **1.3. Model-based optimisation of RLP production**

#### **1.3.1. RLP assembly controllability**

The thermodynamic equilibrium-based model proposed in Chapter 4 provides a means to assess RLP assembly controllability under a particle stability viewpoint. More specifically, the model establishes a relationship between particle formation and the Gibbs free energy of protein subunit association,  $\Delta G_n^0$ , and protein concentration. As discussed before, intracellular protein concentration is a function of the MOI; as such, the MOI is a potential effector of RLP assembly efficiency. The Gibbs free energy of protein subunit association,  $\Delta G_n^0$ , is a very complex function of intracellular physicochemical environment and it is not clear in how far it can be manipulated *in vivo*.

#### Pseudo-critical protein concentration

Model simulations showed that there is a minimum concentration of protein structural subunits needed to ensure a measurable particle formation, i.e., a pseudo-critical concentration, as previously suggested by Zlotnick (1994)



**[24].** At very low concentrations, few particles are formed; above a given critical value, most subunits assemble into intact icosahedra (RLP).

Importantly, this concentration threshold was found to be correlated with the  $\Delta G_n^0$ . From a thermodynamic viewpoint, an increase in the Gibbs free energy of protein subunit association (less negative) means that particle assembly is less favourable. The way to compensate for this loss in intact particles due to  $\Delta G_n^0$  is to increase the concentration of protein subunits in order to shift the equilibrium towards product formation. Thus, at higher  $\Delta G_n^0$ , the protein's pseudo-critical concentration also increases.

#### Maximising RLP assembly efficiency

The model suggests that the key factor for ensuring 100% RLP assembly efficiency is providing vp2, vp6 and vp7 at the exact RLP stoichiometric ratios. Model simulations confirmed that thermodynamics favours the formation of intact icosahedra at low Gibbs free energy values ( $< -2 \text{ kcal.mol}^{-1}$ ) and whenever vp2, vp6 and vp7 are provided at the exact RLP stoichiometric ratios. Significant deviations to this optimal operating point were however observed at high Gibbs free energies, namely a translocation of proteins pseudo-critical concentrations (optimum concentrations are no longer coincident with the stoichiometric ratio of RLP) and a crisp transition from intact icosahedra to unassembled subunits.

#### Gibbs free energy of protein subunit association

From the relationships developed in the thermodynamic equilibrium-based model, we were able to understand and quantify the impact of Gibbs free energy of subunit association on optimal initial vp2, vp6 and vp7 concentrations and on RLP assembly efficiency. A field to be explored in the future is the design of optimization strategies based on oriented manipulation of physico-chemical parameters (e.g. temperature, pH, [NaCl]

and concentration of divalent ions) with impact on Gibbs free energy of subunit association. This is particularly important for the design of *in vitro* assembly/disassembly experiments.

### **1.3.2. RLP optimization by oriented-manipulation of the MOI**

Independently of the infection strategy (single- or co-infection) RLP production and assembly were found to be highly dependent on the MOI (Chapter 6). It should be noted that optimization of RLP production implies the maximisation of protein expression and also the maximisation of assembly efficiency. These two goals are not necessarily compatible and achievable at the same operational MOI.

#### Single infection strategies

For single-infection strategies, model simulations showed that increasing the MOI to very high levels did not bring any advantage as RLP yields are controlled by the stoichiometry of the particle, which is independent of the MOI, and by high protein concentrations, achieved at low MOIs. Indeed, although the intracellular RLP production had increased with the MOI, the total RLP concentration decreased due to lower concentration of infected cells. The assembly efficiency was found to be independent of the MOI.

Interestingly, the profile of RLP production (Figure 38 in Chapter 6) closely resembles the one of vp2 production (Figure 19 in Chapter 3): a minimum at MOIs around 5-10 virus.cell<sup>-1</sup> and a maximum at MOI of 0.01 virus.cell<sup>-1</sup>. These results indicate that optimal MOIs for protein expression in single-infection strategies are independent on the number of recombinant proteins to express.

#### Co-infection strategies

In co-infection strategies, high MOIs maximize RLP production. Since the

concentration of infected cells carrying the genes coding for vp2, vp6 and vp7 was MOI independent (when  $\text{MOI} > 10 \text{ virus.cell}^{-1}$ ) and intracellular RLP production increased with the MOI, the higher the MOI, the higher was the total RLP production.

The model shows that the combination of  $\text{MOI}_2$ ,  $\text{MOI}_6$  and  $\text{MOI}_7$  (number of monocistronic baculoviruses coding for vp2, vp6 and vp7 *per* cell, respectively) maximizing RLP production and assembly efficiency varied considerably. For  $\text{MOI}_2$ ,  $\text{MOI}_6$  and  $\text{MOI}_7$  higher than  $2 \text{ virus.cell}^{-1}$ , these variables are maximized whenever:

- $\text{MOI}_6/\text{MOI}_2 \approx 2.1$
- $\text{MOI}_7/\text{MOI}_2 \approx 1.2$

At these conditions, not only all cells are infected synchronously and by at least one recombinant baculovirus coding for vp2, vp6 and vp7 ( $n_{i,267} \rightarrow n_i$ ) but also the expression levels of the three viral proteins almost obey to the stoichiometry of the particle, thus reducing the amount of waste proteins. At low MOIs ( $\text{MOI}_2$ ,  $\text{MOI}_6$  and  $\text{MOI}_7 < 2 \text{ virus.cell}^{-1}$ ), there is no clear trend in the  $\text{MOI}_6/\text{MOI}_2$  and  $\text{MOI}_7/\text{MOI}_2$  ratios, which is intrinsically related with the high degree of stochasticity of the infection process at such low MOIs.

### 1.3.3. RLP optimization via genetic redesign

#### Genetic controllability

The other alternative is RLP optimization through genetic engineering. A critical hypothesis of the model was that under the control of a given very late promoter (*p10* or *polh*) the transcription and translation rates are linear functions of the gene size. It means that when multiple proteins are expressed under the control of the same promoter, then the relative size of

the gene coding for a given protein determines the relative transcription and translation rates of that particular protein in comparison to others expressed within the same cell. Thus, the rate of protein expression can be controlled either thorough multiple copies of the gene under the control of the same promoter or by the choice of the promoter itself. Although possible in theory, controllability of protein expression at the genetic level may not be straightforward. Noteworthy is the non conformity of the transcriptional rates in co-infection experiments with the rule of the gene size verified in Chapter 5. Although the specific reasons for this discrepancy are not known, one can speculate on the use of different restriction sites for cloning since they have shown to have some influence in the translation and transcription levels [16-21]. For single-infection experiments where the transcription of the viral genes occurs from the same vDNA template and the gene coding for vp7 is under the control of *p10* promoter instead of the *polh* one, the ranking of the transcriptional rates according to gene size was obeyed. This may suggest that the *p10* promoter has a higher potential for vp7 expression than the *polh* promoter. Although inconsistent with literature, where *p10* and *polh* promoters were shown to have similar strengths [22], the levels of gene expression are known to be affected by the interaction between both [23], possibly justifying the aforementioned conclusion.

### Optimisation potential

While in single-infection the redesign of baculovirus vectors at the promoters' level has enormous potential for process optimization, no significant improvements are observed in co-infection strategies (Chapter 6).

In single-infection strategies, model simulations showed that the concentration of infected cells carrying the genes coding for vp2, vp6 and

vp7,  $n_{i,267}$ , was not improved by manipulation of promoters' strength. In contrast, improvements between 7% and 18% in intracellular and total RLP production and around 24% in assembly efficiency were achieved. Irrespective of the MOI evaluated, the combination of promoters' strength maximizing RLP production and assembly efficiency was  $P_{vp6}/P_{vp2} = 1$  and  $P_{vp7}/P_{vp2} = 0.6$  ( $P_{vp2}$ ,  $P_{vp6}$  and  $P_{vp7}$  - the strength of the promoters controlling vp2, vp6 and vp7 expression, respectively).

In co-infection strategies, no significant variation in  $n_{i,267}$  was observed for high MOIs; at lower MOIs, there was an increase in  $n_{i,267}$ . This improvement is more likely to be related with the stochasticity of the infection process than linked to the promoters' strength. This is rational because the promoter strength does not impact directly or indirectly (via vDNA replication rate or synthesis and budding of progeny viruses) on cellular growth; it only affects the transcription and translation mechanisms. Optimal combination of promoters' strengths had a positive impact on intracellular RLP production at high MOIs. By controlling the relative strength of each promoter ( $P_{vp6}/P_{vp2}$  and  $P_{vp7}/P_{vp2}$ ), there was a better adjustment of the expression levels of the three viral proteins to what thermodynamically favours RLP assembly. This reduced the amount of unassembled proteins and increased not only the assembly efficiency but also the total RLP concentration. At low MOIs, no significant variation in intracellular RLP concentration was observed; hence, the increase in total RLP concentration and assembly efficiency is the result of the increase in  $n_{i,267}$ .

### 1.3.4. Single infection or co-infection: what to choose?

In general, single-infection strategies offer significant advantages over co-infection experiments. This is particularly visible at low MOIs where the higher cell density achieved at single-infection led to higher RLP production

(30-60 mg.L<sup>-1</sup> against the 6-12 mg.L<sup>-1</sup> at co-infection). Assembly efficiencies were also higher: 72% against the 20-40% at co-infection. At high MOIs, there was no significant difference between both strategies; the concentration of infected cells carrying the genes coding for vp2, vp6 and vp7, the intracellular and total RLP production and the assembly efficiencies were comparable.

These results are consistent with the ones reported by Belyaev and Roy in 1993 and 1994 for the Bluetongue virus (BTV)-VLP. In their study, the expression of the four main proteins of BTV (vp2, vp3, vp5 and vp7) in a four-gene baculovirus vector has led to the formation of highly stable double-layered BTV-VLPs, contrasting with the single-layered or double-shelled VLPs with highly variable protein ratios obtained by co-infection with two dual-gene baculovirus vectors [13; 14].

## 2. CONCLUSIONS

Based on the work presented in this thesis and the above discussion, the following set of conclusions can be outlined:

- *In silico* baculovirus/insect cells optimization tools are essential for enhancing productivities of complex biotherapeutics in this system. Major steps were undertaken in this thesis towards the *in silico* implementation of baculovirus/insect cells, namely the development of stochastic/thermodynamic-equilibrium models of protein expression and VLP assembly. It combines complex mathematical/statistical formulations (stochasticity, segregation, molecular thermodynamic equilibrium) with experimental intra and extracellular data (vDNA replication, very late transcription, protein synthesis and VLP formation);

- A striking outcome of the model is the independency of vDNA replication rate on the infection strategy. Irrespective of the baculovirus vector used in this study (monocistronic or tricistronic), the sizes of the genes coding for the foreign proteins represent less than 5% of the baculovirus genome size [15]. This explains why the genome size factor had a minor effect on the vDNA replication kinetics;
- The “metabolic burden” effect (decreases the time window for vDNA replication), intracellular vDNA (high vDNA loads result in high mRNA levels) and protein saturation (the translation machinery saturates – no advantage in increasing the MOI to very high levels) are critical factors affecting protein expression;
- The endpoint dilution, microculture tetrazolium and flow cytometric assays were found to be the titration methods combining better cost, time and error. MOI errors can be kept within 20%, which is acceptable for a process optimization program;
- Maximum protein production is achieved at low MOIs, normally in the range of 0.01 to 0.1 pfu.cell<sup>-1</sup> independently of the size of the gene under the control of the *polh* promoter;
- Gene size does not affect the optimal MOI ranges for protein expression but impacts negatively on protein concentrations - the larger the gene size, the lower is the protein concentration;
- Single-infection strategies with tricistronic baculoviruses induce higher RLP production than co-infection strategies with monocistronic baculoviruses;
- The redesign of baculovirus vectors in single-infection strategies, in which the promoters controlling viral protein expression can be selected according to their strength, and the manipulation of the MOI, in both single- and co-infection strategies, have great potential for

process optimization;

- A pseudo-critical protein concentration defines the outcome of the assembly process, either unassembled protein subunits or intact icosahedra. This threshold was found to be correlated with the Gibbs free energy of subunit association,  $\Delta G_n^0$  - as  $\Delta G_n^0$  increases the protein's pseudo-critical concentration also increases;
- RLP yields are thermodynamically maximized whenever vp2, vp6 and vp7 are provided at the exact RLP stoichiometric ratio and at high initial protein concentrations;
- RLP formation can be controlled by oriented-manipulation of solution properties with impact on  $\Delta G_n^0$  (e.g. temperature, pH, ionic strength and divalent ions concentration). This is particularly important in cases where the concentration of proteins is below a pseudo-critical value;
- Inadequate thermodynamic conditions and incorrect protein concentration ratios are responsible for the high percentage of unassembled proteins at the end of a RLP production process.

The present thesis attempts to make the point that in the production of complex biotherapeutics, such as RLP, where bottlenecks are likely to be posed at the molecular level, Systems Biology approaches are essential tools to overcome such bottlenecks. A major benefit is that Systems Biology allows to identify the intrinsic limitations of the production platform and to guide in conformity process development efforts. Moreover, it provides unprecedented level of detail, which in some problems is absolutely essential to find the hidden “secrets” eventually boosting process productivities and product quality. For the model system studied in this thesis, predicted improvements in RLP production would allow increasing the number of vaccine doses produced *per* batch. Indeed, based on recovery yields of 37% upon purification [25] and assuming a vaccine

232



dose of 2-20  $\mu\text{g}$  of RLP in mice [26] or 200-250  $\mu\text{g}$  in gnotobiotic pigs [27; 28], the model-based strategies would provide on average 21 times more vaccine doses than traditional RLP production processes ( $\sim 2 \text{ mg/L}$ ). Even if the RLP yields calculated in Chapter 6, based on particle stoichiometry and the limiting protein, had been calculated based on literature assembly efficiencies ( $\sim 11\%$  [29]), a 3-fold increase would still be expected.

These theoretical results suggest that the integration of the model-based strategies explored in this thesis within a flexible vaccine production platform will be essential to accelerate the timeline for moving RLP vaccine candidate into the market.

### 3. FUTURE WORK

This thesis represents a contribution for the clarification of some critical aspects of RLP production in the baculovirus/insect cell system and for the design of optimization schemes capable of maximizing the objective function, RLP production. In addition, it provides a computational framework with potential for process production development and optimization of viral proteins and virus-like particles in a broad sense.

Still, to achieve a robust and efficient RLP production platform, some trends identified throughout this thesis need verification:

- The positive impact on RLP production of redesigning baculovirus vectors, in which the promoters controlling viral protein expression can be selected according to their strength;
- The increase in RLP production using the optimal combination of  $\text{MOI}_2$ ,  $\text{MOI}_6$  and  $\text{MOI}_7$  at co-infection and the best MOI at single-infection;

- Control the formation of RLPs *in vivo* by oriented manipulation of solution properties with impact on Gibbs free energy of subunit association (e.g. temperature, pH, ionic strength and divalent ions concentration); this is even more essential if an assembly/reassembly step is required for improved purification.

whereas other issues require further investigation:

- Evaluate the effect of promoters' class (early, late or very late promoter), TOI and CCI on the performance of single or co-infection strategies;
- Integration of the information coming from -omic technologies into a global optimization model for supporting decision-making at the engineering level;
- Development of optimization strategies based on metabolic engineering;
- The impact of different culture modes (e.g. perfusion) and key process parameters (e.g. temperature shifts) on RLP production;
- From a clinical perspective - the formulation of the vaccine (buffer, adjuvant or delivery system), the dosage and the storage condition;
- Generation of chimeric RLPs for broader protection against different rotavirus serotypes or for the development of newer vaccines against emergent diseases;
- Application of this modelling framework to other economically relevant viruses (e.g. adeno-associated virus), heterologous proteins or VLPs.

## 4. REFERENCES

1. FDA, **1998**. Guidance for Industry Q5A Viral Safety Evaluation of Biotechnology Products Derived From Cell Lines of Human or Animal Origin. 27.
2. Mena, J., Ramirez, O.T., Palomares, L., **2003**. Titration of Non-Occcluded Baculovirus Using a Cell Viability Assay. *Biotechniques*. 34 (2): 260-264.
3. LaBarre, D.D., Lowy, R.J., **2001**. Improvements in methods for calculating virus titer estimates from TCID50 and plaque assays. *J Virol Methods*. 96 (2): 107-126.
4. Kompier, R., Tramper, J., Vlask, J.M., **1988**. A continuous process for the production of baculovirus using insect cell cultures. *Biotechnol. Lett.* 10 849-854.
5. Gotoh, T., Chiba, K., Kikuchi, K., **2004**. Probabilistic Characterization for Baculovirus-Infected Insect Cells Destined to Synthesize Progeny Viruses and Recombinant Protein and to Die. *J Chem Eng Jpn*. 37 (11): 1357-1366.
6. Hu, Y.-C., Bentley, W.E., **2000**. A kinetic and statistical-thermodynamic model for baculovirus infection and virus-like particle assembly in suspended insect cells. *Chem Eng Sci*. 55 3991-4008.
7. Maranga, L., Brazao, T.F., Carrondo, M.J., **2003**. Virus-like particle production at low multiplicities of infection with the baculovirus insect cell system. *Biotechnol Bioeng*. 84 (2): 245-53.
8. Power, J.F., Reid, S., Radford, K.M., Greenfield, P.F., Nielsen, L.K., **1994**. Modeling and optimization of the baculovirus expression vector system in batch suspension culture. *Biotechnol Bioeng*. 44 (6): 710-719.
9. Power, J.F., Nielsen, L.K., **1996**. Modelling baculovirus infection of insect cells in culture. *Cytotechnology*. 20 (1-3): 209-219.
10. Licari, P., Bailey, J.E., **1992**. Modeling the population dynamics of baculovirus-infected insect cells: Optimizing infection strategies for enhanced recombinant protein yields. *Biotechnol Bioeng*. 39 (4): 432-441.
11. Rosinski, M., Reid, S., Nielsen, L.K., **2002**. Kinetics of baculovirus replication and release using real-time quantitative polymerase chain reaction. *Biotechnol Bioeng*. 77 (4): 476-480.
12. Hu, Y.C., Bentley, W.E., **2001**. Effect of MOI ratio on the composition and yield of chimeric infectious bursal disease virus-like particles by baculovirus co-infection: deterministic predictions and experimental results. *Biotechniques*. 75 (1): 104-119.
13. Belyaev, A.S., Roy, P., **1993**. Development of baculovirus triple and quadruple expression vectors: co-expression of three or four bluetongue virus proteins and the synthesis of bluetongue virus-like particles in insect cells. *Nucleic Acids Res*. 21 (5): 1219-23.
14. Roy, P., **1994**. Use of baculovirus expression vectors to study protein-protein interactions in bluetongue virus assembly. *Molecular Biology*. 13 183-192.
15. Summers, M.D., Smith, G.E., **1978**. Baculovirus structural polypeptides. *Virology*. 84 (2): 390-402.
16. van Oers, M.M., Vlask, J.M., Voorma, H.O., Thomas, A.A., **1999**. Role of the 3' untranslated region of baculovirus p10 mRNA in high-level expression of foreign genes. *J Gen Virol*. 80 (Pt 8): 2253-2262.
17. Qin, J.C., Liu, A.F., Weaver, R.F., **1989**. Studies on the control region of the p10 gene of the Autographa californica nuclear polyhedrosis virus. *J Gen Virol*. 70 (Pt 5): 1273-1279.
18. Ooi, B.G., Rankin, C., Miller, L.K., **1989**. Downstream sequences augment transcription from the essential initiation site of a baculovirus polyhedrin gene. *J Mol Biol*. 210 (4): 721-736.
19. Weyer, U., Possee, R.D., **1988**. Functional analysis of the p10 gene 5' leader sequence of the Autographa californica nuclear polyhedrosis virus. *Nucleic Acids Res*. 16 (9): 3635-3653.
20. Luckow, V.A., Summers, M.D., **1988**. Signals important for high-level expression of foreign genes in Autographa californica nuclear polyhedrosis virus expression vectors. *Virology*. 167 (1): 56-71.
21. Possee, R.D., Howard, S.C., **1987**. Analysis of the polyhedrin gene promoter of the Autographa californica nuclear polyhedrosis virus. *Nucleic Acids Res*. 15 (24): 10233-10248.
22. van Oers, M.M., Malame, D., Jore, J.M., Vlask, J.M., **1992**. Expression of the Autographa californica nuclear polyhedrosis virus p10 gene: effect of polyhedrin gene expression. *Arch Virol*. 123 (1-2): 1-11.
23. Chaabihi, H., Ogliastro, M.H., Martin, M., Giraud, C., Devauchelle, G., Cerutti, M., **1993**. Competition between baculovirus polyhedrin and p10 gene expression during infection of insect cells. *J Virol*. 67 (5): 2664-2671.
24. Zlotnick, A., **1994**. To build a virus capsid. An equilibrium model of the self assembly of polyhedral protein complexes. *J Mol Biol*. 241 (1): 59-67.
25. Peixoto, C., Sousa, M.F., Silva, A.C., Carrondo, M.J., Alves, P.M., **2007**. Downstream processing of triple layered rotavirus like particles. *J Biotechnol*. 127 (3): 452-61.

26. Istrate, C., Hinkula, J., Charpilienne, A., Poncet, D., Cohen, J., Svensson, L., Johansen, K., **2008**. Parenteral administration of RF 8-2/6/7 rotavirus-like particles in a one-dose regimen induce protective immunity in mice. *Vaccine*. 26 (35): 4594-601.
27. Gonzalez, A.M., Nguyen, T.V., Azevedo, M.S., Jeong, K., Agarib, F., Iosef, C., Chang, K., Lovgren-Bengtsson, K., Morein, B., Saif, L.J., **2004**. Antibody responses to human rotavirus (HRV) in gnotobiotic pigs following a new prime/boost vaccine strategy using oral attenuated HRV priming and intranasal VP2/6 rotavirus-like particle (VLP) boosting with ISCOM. *Clin Exp Immunol*. 135 (3): 361-72.
28. El-Attar, L., Oliver, S.L., Mackie, A., Charpilienne, A., Poncet, D., Cohen, J., Bridger, J.C., **2009**. Comparison of the efficacy of rotavirus VLP vaccines to a live homologous rotavirus vaccine in a pig model of rotavirus disease. *Vaccine*. 27 (24): 3201-8.
29. Vieira, H., Estêvão, C., Roldão, A., Peixoto, C., Sousa, M., Cruz, P.E., Carrondo, M.J.T., Alves, P.M., **2005**. Triple layered rotavirus VLP assembly: kinetics of vector replication, mRNA stability and recombinant protein production. *J Biotechnol*. 120 (1): 72-82.