

Novel systems biology approaches for design and analysis of palmitoylation networks reveal its regulatory effects on protein stability, activity and localization

THÈSE N° 7383 (2017)

PRÉSENTÉE LE 1^{ER} MARS 2017

À LA FACULTÉ DES SCIENCES DE BASE

LABORATOIRE DE BIOTECHNOLOGIE COMPUTATIONNELLE DES SYSTÈMES
PROGRAMME DOCTORAL EN BIOTECHNOLOGIE ET GÉNIE BIOLOGIQUE

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Tiziano DALLAVILLA

acceptée sur proposition du jury:

Prof. M. Dal Peraro, président du jury
Prof. V. Hatzimanikatis, Prof. F. G. van der Goot Grunberg, directeurs de thèse
Prof. L. Chamberlain, rapporteur
Prof. M. Barberis, rapporteur
Prof. O. Hantschel, rapporteur



ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2017

Abstract

Communication and environment sensing are fundamental for the regulation of the majority of the cellular functions. In cells there are dedicated networks that are responsible for signalling, which gives the ability to the cell of sensing the external and internal environment allowing rapid response and adaptation to changes and stress events. A major component of these networks are post translational modifications. Their attachment to proteins in response to a change in the surrounding conditions can have very drastic consequences on modified proteins. Since post translational modification can heavily influence the cellular fate it is essential to identify these modifications and to understand their functioning, to better comprehend how cells respond to environments changes and different types of stress events to which they are continuously subjected.

In this work we analyse a type of modification called S-palmitoylation, which is recently emerging as signalling/regulatory modification. S-Palmitoylation, or more generally “acylation”, is the addition of an acyl chain to the SH-group of a cysteine via a thioester bond. This type of modification has a big impact at the cellular level. In different studies about palmitoylated proteins, functional consequences are observed at the cellular level, like altered signalling capacity, activity modulation, regulation of protein stability and localization.

Palmitoylation was studied using a mathematical modelling approach. Based on experimental data we developed models of the palmitoylation process of different proteins. Using a bottom up approach we first investigated the consequences of protein palmitoylation on the endoplasmic reticulum chaperone calnexin, to better understand the consequences of palmitoylation at the protein level. Following the same approach we investigated the effects of palmitoylation on DHHC6, a

palmitoyltransferase that is responsible for palmitoylation of calnexin, which itself undergoes palmitoylation from the palmitoyltransferase DHHCl6. In the last part of the project we focused on the study of palmitoylation at the network level. For this purpose we reconstructed the palmitoylation cascade of calnexin, including in the model all the proteins involved in the palmitoylation process of this protein. This network of palmitoylation allowed us to investigate the dynamics of palmitoylation at the network level, focusing in the characterization of those mechanisms that are responsible for regulation of palmitoylation on different substrates.

This work provides an unprecedented level of understanding of palmitoylation dynamics, both at the protein and network level. Moreover all the models and the tools for model analysis used in this work have been developed keeping in mind the final goal of reconstructing the entire palmitoylation network. Thanks to the approach we chose for modelling, the calnexin network model developed during this project can be easily expanded to include other portions of the palmitoylation networks. Similarly the tools developed for model analysis can be easily adapted to work on post-translational modification assays and analysis of biological functions from the smaller to the larger scales.

Keywords: palmitoylation, signalling, post translational modification, systems biology, network, acylation, regulation.

Sommario

Il rilevamento delle condizioni ambientali e la comunicazione sono due aspetti fondamentali per la regolazione della maggioranza delle funzioni cellulari. Nelle cellule esistono network responsabili per l'interpretazione dei segnali ambientali, che rendono la cellula capace di percepire le condizioni interne ed esterne, conferendo la capacità di adattarsi rapidamente ai cambiamenti e allo stress. Una componente importante di questi network sono le modificazioni post traduzionali. La loro aggiunta alle proteine in risposta ad un cambiamento delle condizioni ambientali può avere conseguenze drastiche sulle proteine modificate. Siccome le modificazioni post traduzionali possono avere pesanti conseguenze sul destino di una cellula è essenziale identificare queste modificazioni e comprendere il loro funzionamento, per capire meglio come le cellule rispondono ai cambiamenti ambientali e ai differenti tipi di stress alle quali sono continuamente soggette.

In questo lavoro analizziamo un tipo di modificazione chiamata S-palmitoilazione che recentemente sta emergendo come modificazione regolatoria/di segnalazione. La S-palmitoilazione, o più generalmente "acilazione", consiste nell'aggiunta di una catena acilica al gruppo SH di una cisteina attraverso un legame tioestere. Questo tipo di modificazione ha un grande impatto a livello cellulare, diversi studi riguardanti la palmitoilazione di proteine hanno osservato diverse conseguenze funzionali a livello cellulare, come ad esempio un'alterata capacità di segnalazione, di modulazione dell'attività enzimatica e della regolazione della stabilità e localizzazione proteica.

In questo progetto la palmitoilazione è stata studiata utilizzando un approccio di modellazione matematica. Sulla base di dati sperimentali abbiamo sviluppato modelli del processo di palmitoilazione di diverse proteine. Utilizzando un approccio bottom up abbiamo investigato le conseguenze della palmitoilazione della calnexina, un chaperone molecolare residente nel reticolo endoplasmatico, al fine di comprendere meglio le conseguenze della a livello di proteina. Seguendo un approccio simile

abbiamo investigato gli effetti della palmitoilazione su DHHC6, una palmitoiltransferasi responsabile della palmitoilazione della calnexina, che è essa stessa palmitoilata dalla palmitoiltransferasi DHHC16. Nell'ultima parte del progetto ci siamo focalizzati sullo studio della palmitoilazione a livello sistemico. A questo scopo abbiamo ricostruito la cascata di palmitoilazione della calnexina, includendo nel modello tutte le proteine coinvolte nel processo di palmitoilazione di questa proteina. Questo network ci ha permesso di investigare le dinamiche di palmitoilazione a livello sistemico focalizzandoci principalmente nella caratterizzazione di quei meccanismi responsabili della regolazione della palmitoilazione di diversi substrati.

Questo lavoro fornisce un livello di comprensione delle dinamiche di palmitoilazione senza precedenti, sia a livello di proteina che di network. Oltretutto, tutti i modelli e gli strumenti utilizzati in questo progetto sono stati sviluppati in funzione dell'obiettivo finale di ricostruzione dell'intero network cellulare di palmitoilazione. Grazie all'approccio scelto per la modellazione, il modello descrittore della cascata di palmitoilazione della calnexina può essere facilmente espanso al fine di includere altre porzioni del network di palmitoilazione. Allo stesso modo gli strumenti sviluppati per l'analisi dei modelli possono essere facilmente adattati per lavorare con diversi tipi di dati sperimentali e per l'analisi di modelli di diversa scala, dalla singola proteina all'intero network cellulare di palmitoilazione.

Parole chiave: palmitoilazione, signalling, modificazioni post traduzionali, systems biology, network, acilazione, signalling, regolazione.

Table of Contents:

ABSTRACT	1
CHAPTER 1 - INTRODUCTION	12
CHAPTER 2 - STATE OF THE ART	17
2.1 WHAT IS PALMITOYLATION?	17
2.2 PALMITOYLATION IN HUMANS	19
2.3 THE ROLE OF PALMITOYLATION IN THE REGULATION OF CELLULAR FUNCTIONS.....	20
2.4 PALMITOYLATION CYCLE: DHHCS AND APTS	22
2.5 STUDY PALMITOYLATION WITH THE HELP OF A MODELLING APPROACH.....	25
2.1 STUDY PALMITOYLATION WITH THE HELP OF MODELLING.....	26
CHAPTER 3 - MATERIAL AND METHODS.....	30
3.1 MODELING OF PALMITOYLATION NETWORKS: DEFINE THE NETWORK TOPOLOGY	31
3.2 APPLICATION OF TQSSA TO THE ENZYMATIC REACTIONS OF PALMITOYLATION	33
3.3 DEVELOPMENT OF AN IN-SILICO LABELLING TECHNIQUE FOR MODEL CALIBRATION AND ANALYSIS	42
3.3.1 ³⁵ S labelling.....	44
3.3.2 ³ H labeling	48
3.3.3 Model parameterization with heuristic optimization	49
3.3.4 Studying impact of modelling parameters with sensitivity analysis.....	53
3.4 STOCHASTIC SIMULATION AS A TOOL FOR MOLECULE TRACKING.....	59
3.5 STOCHASTIC MODEL FORMULATION:	61
3.6 CONVERSION OF DETERMINISTIC PARAMETERS TO STOCHASTIC:	62
CHAPTER 4 - MODEL-DRIVEN UNDERSTANDING OF PALMITOYLATION DYNAMICS: REGULATED ACYLATION OF THE ENDOPLASMIC RETICULUM CHAPERONE CALNEXIN.....	64
4.1 ABSTRACT.....	64
4.2 AUTHOR SUMMARY.....	65
4.3 INTRODUCTION.....	65
4.4 RESULTS.....	68
4.4.1 Mathematical model of calnexin palmitoylation.....	68
4.4.2 Experimental analysis of the effect of palmitoylation on calnexin stability	71
4.4.3 Experimental determination of palmitate turnover on Calnexin.....	72
4.4.4 Calibration and validation of the mathematical model.....	74
4.4.5 Predicted species distribution of calnexin.....	75

4.4.6	<i>Estimated half-lives of palmitoylated calnexin species</i>	79
4.4.7	<i>Cooperativity between the two calnexin palmitoylation sites</i>	81
4.4.8	<i>Efficient degradation of calnexin is predicted to require depalmitoylation</i>	82
4.4.9	<i>Estimation of the lag-time separating calnexin synthesis from palmitoylation</i>	82
4.4.10	<i>Phosphorylation is a negative regulator of calnexin palmitoylation</i>	83
4.5	DISCUSSION	85
4.6	MATERIAL AND METHODS	88
4.7	SUPPLEMENTARY MATERIAL	91
4.8	APPENDIX A	96
CHAPTER 5 -	DHHC16-MEDIATED CONTROL OF DHHC6 ACTIVITY: DYNAMICS OF A	
PALMITOYLATION CASCADE		118
5.1	INTRODUCTION	118
5.2	RESULTS	121
5.2.1	<i>Palmitoylation of the DHHC6 SH3 domain</i>	121
5.2.2	<i>The DHHC16-DHHC6 palmitoylation cascade</i>	122
5.2.3	<i>Rapid APT2-mediated DHHC6 depalmitoylation</i>	124
5.2.4	<i>Palmitoylation of Cys-328 destabilizes DHHC6</i>	125
5.2.5	<i>Palmitoylation-dependent DHHC6 localization</i>	126
5.2.6	<i>Palmitoylation dependent activity of DHHC6</i>	128
5.2.7	<i>Model of the DHHC6 palmitoylation system</i>	129
5.2.8	<i>Dynamics of DHHC6 palmitoylation and effect on stability</i>	131
5.2.9	<i>Importance of multiple palmitoylation sites for the regulation of DHHC6 activity</i>	133
5.3	DISCUSSION	136
5.4	MATERIAL AND METHODS	141
5.5	APPENDIX B	150
CHAPTER 6 -	RECONSTRUCTION OF A PALMITOYLATION NETWORK	175
6.1	INTRODUCTION	175
6.2	NETWORK DESIGN	175
6.3	PRELIMINARY RESULTS	180
6.3.1	<i>Model calibration</i>	180
6.3.2	<i>Global sensitivity analysis</i>	182
6.4	CONCLUSION	187
	APPENDIX C	189
CHAPTER 7 -	FUTURE WORK	196

CHAPTER 8 -	CONCLUSIONS	201
CHAPTER 9 -	REFERENCES	206

List of Tables

TABLE 4.1. HALF-LIFE OF THE DIFFERENT CALNEXIN SPECIES.....	81
TABLE S4.2. MODEL REACTIONS	105
TABLE S4.3. MODEL PARAMETERS.	109
TABLE S4.4. MASS BALANCE EQUATIONS.....	110
TABLE S4.5. STOICHIOMETRIC MATRIX USED FOR STOCHASTIC SIMULATIONS.....	111
TABLE 4.6. PROPENSITY FUNCTION USED FOR STOCHASTIC SIMULATIONS.	112
TABLE S4.7. PARAMETERS USED FOR STOCHASTIC SIMULATIONS.....	117
TABLE S5.1 HALF-LIFE OF DHHC6 IN DIFFERENT PALMITOYLATION STATES.	156
TABLE S5.2. TOTAL AMOUNT OF PROTEIN IN STEADY STATE RELATIVE TO WT FOR WT AND CAA MUTANT IN DIFFERENT CONDITIONS.	156
TABLE S5.3 RESULTS OF STOCHASTIC SIMULATION WHEN APT2 IS SILENCED.....	157
TABLE S5.4. MODEL REACTIONS.	157
TABLE S5.5. MODEL PARAMETERS.	163
TABLE S5.6. MASS BALANCE EQUATIONS.....	165
TABLE S5.7.. PARAMETERS USED FOR STOCHASTIC SIMULATIONS.....	166
TABLE S5.8 STOICHIOMETRIC MATRIX USED FOR STOCHASTIC SIMULATIONS.....	168
TABLE 5.9. PROPENSITY FUNCTION USED FOR STOCHASTIC SIMULATIONS.	169
TABLE 6.1. RESULTS OF PARAMETER ESTIMATION.	189
TABLE 6.2. GROUPS USED FOR GLOBAL SENSITIVITY ANALYSIS	192

List of figures

FIGURE 2.1. REACTION MECHANISM OF PALMITOYLATION.....	18
FIGURE 2.2. GENERAL STRUCTURE OF A DHHC ENZYME.....	23
FIGURE 3.1 GENERIC PROTEIN PALMITOYLATION/DEPALMITOYLATION CYCLE.....	31
FIGURE 3.2 MODEL OF CALNEXIN PALMITOYLATION/DEPALMITOYLATION CYCLE.....	32
FIGURE 3.3. MASS ACTION MODEL OF CALNEXIN PALMITOYLATION.....	38
FIGURE 3.4. NETWORK TOPOLOGY OF THE CALNEXIN MODEL USED FOR ³⁵ S <i>IN-SILICO</i> LABELLING..	ERROR! BOOKMARK NOT DEFINED.
FIGURE 3.5. DIFFERENT PHASES OF A ³⁵ S <i>IN-SILICO</i> LABELLING EXPERIMENT.....	47
FIGURE 3.6. NETWORK TOPOLOGY OF THE CALNEXIN MODEL USED FOR ³ H <i>IN-SILICO</i> LABELLING.....	48
FIGURE 3.7. DIFFERENT PHASES OF A ³ H <i>IN-SILICO</i> LABELLING EXPERIMENT.....	49
FIGURE 3.8. DIFFERENT PHASES OF A SINGLE MOLECULE TRACKING EXPERIMENT.....	60
FIGURE 3.9. RESULTS OF MODEL CALIBRATION AND VALIDATION FOR A SINGLE SET OF PARAMETERS OBTAINED BY AVERAGING GA OUTPUT.....	63
FIGURE 4.1. CALNEXIN PALMITOYLATION MODEL AND KINETICS OF DECAY.....	69
FIGURE 4.2. EXPERIMENTS AND MODELLING OF PALMITOYLATION/DEPALMITOYLATION OF CALNEXIN.....	73
FIGURE 4.3. <i>IN SILICO</i> ANALYSIS OF THE CALNEXIN SPECIES DISTRIBUTION.....	75
FIGURE 4.4. PREDICTION OF DEPALMITOYLATION KINETICS.....	78
FIGURE 4.5. KINETICS OF CALNEXIN PALMITOYLATION AND STABILITY.....	79
FIGURE 4.6. STOCHASTIC SIMULATIONS REVEAL THE AVERAGE PALMITOYLATION TIME FOR CALNEXIN.....	83
FIGURE 4.7. PREMATURE CALNEXIN PALMITOYLATION IS PREVENTED BY SERINE PHOSPHORYLATION.....	84
FIGURE S4.8. RESULTS OF THE GA OPTIMIZATION ARE PLOTTED ON TOP OF THE EXPERIMENTAL DATA USED AS OBJECTIVES.....	96
FIGURE S4.9. SCHEMATIC REPRESENTATION OF THE MODEL AND THE PROCEDURE USED FOR ³⁵ S PULSE-CHASE EXPERIMENTS SIMULATION.....	97
FIGURE S4.10. SCHEMATIC REPRESENTATION OF THE MODEL AND THE PROCEDURE USED FOR ³ H PULSE-CHASE EXPERIMENTS SIMULATION.....	98
FIGURE S4.11. DISTRIBUTION OF THE LABELLED SPECIES IN THE DIFFERENT PALMITOYLATION STATES AFTER 2 HOURS LABELING WITH ³ H PALMITATE.....	99
FIGURE S4.12. QUALITATIVE REPRESENTATION OF THE PROCEDURE USED FOR PROTEIN TRACKING WITH STOCHASTIC SIMULATIONS, ALONG WITH PROOF OF CONVERGENCE.....	100
FIGURE S4.13. STEADY STATE SENSITIVITY ANALYSIS OF CALNEXIN WT AND MUTANTS.....	101
FIGURE S4.14. HALF LIFE SENSITIVITY ANALYSIS OF CALNEXIN WT.....	102
FIGURE S4.15. HALF LIFE SENSITIVITY ANALYSIS OF CALNEXIN AA MUTANT.....	102
FIGURE S4.16. HALF LIFE SENSITIVITY ANALYSIS OF CALNEXIN AC MUTANT.....	103
FIGURE S4.17. HALF LIFE SENSITIVITY ANALYSIS OF CALNEXIN CA MUTANT.....	103
FIGURE S4.18. FITTING OF THE DATA USED FOR CALIBRATION AND VALIDATION AFTER AVERAGING THE 383 SETS OF PARAMETERS ESTIMATED BY THE GA.....	104

FIGURE 5.1. DHHC6 CAN UNDERGO PALMITOYLATION ON THREE CYSTEINS OF ITS SH3 DOA.....	120
FIGURE 5.2. DHHC6 IS PALMITOYLATED BY DHHC16.....	123
FIGURE 5.3. A. DEGRADATION KINETICS OF DHHC6.....	124
FIGURE 5.4.....	127
FIGURE 5.5. NETWORK TOPOLOGY OF THE DHHC6 PALMITOYLATION MODEL.....	129
FIGURE 5.6. RESULTS OF SINGLE MOLECULE TRACKING WITH STOCHASTIC SIMULATIONS WHEN DHHC16 IS OVEREXPRESSED..	134
FIGURE S5.7 ALIGNMENT OF DHHC6 SEQUENCES FROM DIFFERENT SPECIES. IDENTICAL AMINO ACIDS ARE HIGHLIGHTED BY DARK BLUE.....	150
FIGURE S5.8. DHHC6 IS NOT AUTOPALMITOYLATED.....	151
FIGURE S5.9. CO-IMMUNOPRECIPITATION OF DHHC6 WITH DHHC16.....	152
FIGURE 5.10 LEVEL OF MUTANT DHHC6 PALMITOYLATION.....	153
FIGURE S5.11 CALIBRATION.....	154
FIGURE S5.12. VALIDATION.....	155
FIGURE 6.1. REPRESENTATION OF CALNEXIN PALMITOYLATION CASCADE.....	176
FIGURE 6.2. ILLUSTRATION OF THE COMPLEXITY RISING FROM THE PRESENCE OF MULTIPLE MODIFICATION SITES ON PROTEIN AND THE IMPLEMENTATION OF <i>IN-SILICO</i> LABELLING TECHNIQUES.....	178
FIGURE 6.3. GLOBAL SENSITIVITY ANALYSIS RESULTS FOR TOTAL DHHC6/CALNEXIN IN STEADY STATE.....	183
FIGURE 6.4. GLOBAL SENSITIVITY ANALYSIS ON PALMITATE INCORPORATION RATE OF DHHC6 AND CALNEXIN.....	186

Abbreviations:

APT: Acyl Protein Thioesterase

ER: Endoplasmic reticulum

GA: Genetic algorithm

MM: Michaelis-Menten

ODE: Ordinary differential equation

PAT: Palmitoyltransferase

P.E.: Parameter estimation

PTM: Post-translational modification

QSSA: Quasi-steady state approximation

SA: Sensitivity analysis

tQSSA: Total quasi-steady state approximation

Chapter 1 - Introduction

Cellular communication is fundamental for the regulation of the majority of cellular functions. Signalling gives the cell the ability to sense the external and internal environment, allowing rapid response and adaptation to changes and stress events. One of the most important component of the signalling networks are post translational modifications (PTMs), which are additions of different chemical groups to the amino acid residues of proteins. PTMs are the carriers of signals generated from environment sensing. Their attachment to proteins in response to a change in the surrounding conditions can have drastic consequences on the modified peptides and, in turn, on their downstream cellular functions.

There are many chemical groups that can be attached to the side chain of various types of amino acids, and each one has different chemical properties and consequences [1]. Different proteins rely on modification of one or more amino acid residues for their activation/deactivation. PTMs also have a central role in the intracellular localization of proteins [2]. Additionally, several type of PTMs have been implicated in initiation of conformational changes [1]. Among the different PTMs the most interesting are those that participate in the cell signalling/regulatory systems, because of their role in regulation of cellular functions. Two of the most studied PTMs that are part of signalling/regulatory networks are phosphorylation [3] and ubiquitination [4]. Phosphorylation consists in the addition of a phosphate group (PO_4^{3-}) to a serine, threonine or tyrosine residue of a protein. The first evidence of this modification was found at the beginning of the 20th century, but its characterization as an enzymatic reaction occurred only much later, in 1954 [5]. Phosphorylation became popular very quickly among the scientific community and today is still one of

the main areas of research, with more than 240,000 articles on the PubMed database. One of the prominent characteristics of this PTM is that its attachment is reversible, meaning the phosphate group can be attached on a substrate by a protein kinase, and it can be removed later on by a phosphatase. Reversibility allows modifying a protein temporarily and only when it is needed, a fundamental step in the formation of complex regulatory networks. Another reason for its popularity is that phosphorylation is the most widespread modification; it is estimated that 30% of the entire human proteome is phosphorylated at least on one residue [6,7]. Abnormal functioning of phosphorylation is associated with rise of different pathologies such as Alzheimer [8], Parkinson [9] and cancer [10].

Ubiquitination is another signalling post translational modification. It was discovered in 1975 by Gideon Goldstein [11]. One of the most interesting characteristics is that the molecule that modifies the proteins is itself a 8.5 kDa peptide called ubiquitin. Just like phosphorylation, ubiquitination is reversible. Moreover, multiple ubiquitins can be attached to the same lysine to form polyubiquitin chains. The enzymatic reaction of ubiquitination is a quite complex process involving three different enzymes, a ubiquitin-activating enzyme (E1), a ubiquitin-conjugating enzyme (E2) and a ubiquitin ligases (E3). The opposite reaction of deubiquitinating is catalysed by a class of enzymes called deubiquitinating enzymes. They are cysteine proteases that cleave the amide bonds between ubiquitin and the modified proteins. This modification can have various effects on proteins, such as polyubiquitination promoting proteins degradation [12]. It can also affect protein localization and can promote, or prevent, proteins interactions [13,14].

Recently a new type of modification is emerging as a signalling/regulatory molecule -- S-Palmitoylation. It is gaining importance since it shows similar properties to phosphorylation and ubiquitination, suggesting that this modification may be another central part of different cellular regulatory networks. Palmitoylation belongs to the class of PTMs called "acylation", consisting in the addition of fatty acid molecules to the side chain of an amino acid residue. For palmitoylation the fatty acid

attached to the protein is a saturated sixteen or eighteen carbon atom molecules ($\text{CH}_3\text{-(CH}_2\text{)}_{14}\text{-CO}$ or $\text{CH}_3\text{-(CH}_2\text{)}_{16}\text{-CO}$) called palmitate and stearic acid respectively. The modification is attached to the residue at the SH-group of a cysteine via thioester bond [15–17]. To date, hundreds of proteins have already been found to be palmitoylated in mammals [18], and more than thousand proteins are predicted to be palmitoylated based on Uniprot annotations [19]. It is believed, based on predictions, that more than 12% of the human proteome may undergo palmitoylation [16]. Just as in phosphorylation and ubiquitination, this modification is reversible, and remarkably this is the only lipid modification showing this property. Also like phosphorylation and ubiquitination, this property is often observed in those modifications that have a major role in the signalling network. Unlike phosphorylation, which has been extensively studied for more than fifty years, palmitoylation has attracted the interest of the scientific community only recently. The main reason is that even though palmitoylation was discovered during the early fifties, the identification of the enzymes catalysing this PTM has been very difficult, mainly because these enzymes are membrane bound, therefore more difficult to purify and analyse. As a consequence, today the role of palmitoylation, its regulation, its dynamics and consequences are much less understood [15,20]. As experimental techniques advance and new approaches for network analysis are being developed, an increasing number of studies are uncovering its role as a signalling molecule in the cell, showing that palmitoylation is involved in regulation of many fundamentals cellular functions. In various studies (for a review see [15,17,20]) about palmitoylated proteins, functional consequences are observed at the cellular system level, some of the most intriguing being altered signalling capacity [21–23] and activity modulation [24].

In this thesis project, we focused on the study of palmitoylation and its consequences at the protein and network level. In Chapter 2 - we review the state of the art about palmitoylation. We initially define what palmitoylation is and we introduce the mechanism of modification, talking about the enzymes that are responsible for the cycle of palmitoylation/depalmitoylation. We then show how extended this network is and how palmitoylation is capable of interacting with other post-translational

modifications in the cell. Finally, we give an overview about the known functions and consequences at the protein and system level.

To study this PTM, we used a mathematical modelling approach: based on experimental data we developed models of the palmitoylation process of different modified proteins. This approach implied the development of different tools for model design and analysis which are described in Chapter 3 - . First we introduce kinetic modelling as a tool for signalling and regulatory models design, and we show how this was used to define the reaction mechanism of a palmitoylation event. We then introduce the techniques that we applied to perform model reduction, in order to decrease the dimensionality of the system, its complexity and the computation time. In the second part of this chapter we talk about the tools used for model analysis. We introduce an *in-silico* labelling technique developed *ad-hoc* for parameter estimation from radio-labelling experiments *in-vivo*. In this part, we also describe the procedure and the algorithm used for parameter estimation, along with the local and global algorithms used to perform sensitivity analysis. The final section is dedicated to the development of a method allowing to perform single molecule tracking using stochastic simulations. To our knowledge, this is the first method that allows molecule tracking using stochastic simulations.

In Chapter 4 - we illustrate the development and analysis of the palmitoylation model of the endoplasmic reticulum (ER) chaperone calnexin. Using a bottom-up approach we describe the double step process of calnexin palmitoylation, and we show how the model was investigated to improve the understanding of the dynamics and the consequences of palmitoylation on calnexin. Using a similar approach in Chapter 5 - we investigate the dynamics of palmitoylation of the palmitoyltransferase DHHC6. DHHC6 is an enzyme that catalyses palmitoylation and is itself palmitoylated on three different sites. This raised the intriguing possibility that palmitoylation enzymes form a complex regulatory network that tightly regulates palmitoylation. Although the approach used to develop DHHC6 model is similar to what was done for calnexin, its analysis presented some new challenges rising from

the increased complexity of the palmitoylation process. This is partly due to the increased number of palmitoylation sites, and the dynamic nature of the modification on this enzyme.

In the last part of the project the different single protein models were used as starting point for the reconstruction of a small palmitoylation sub network. In Chapter 6 - we analyse the palmitoylation network of calnexin that was generated starting from the single protein models of calnexin and DHHC6. The network was developed using rule-based modelling, which allows to define rules of interaction that automatically generate the kinetic model of ordinary differential equations (ODEs). The focus of this section was to investigate the properties of palmitoylation at the system level, studying those features that can't be captured by single protein models. The network model improved our understanding of calnexin palmitoylation process; in particular, it showed how it is possible to correlate the dynamics of palmitoylation of calnexin with those of DHHC6, demonstrating how the level of palmitoylation and the abundance of calnexin are governed by DHHC6 and other upstream DHHC enzyme involved in the palmitoylation cascade of calnexin. One of the best features of the network model is that it can be generalized to represent and study other portions of the network that doesn't belong to the calnexin palmitoylation system, since many different palmitoylated proteins undergo a palmitoylation cascade similar to what is observed for calnexin.

The models of calnexin along with the tools for its analysis have been already used in 1 publication, while DHHC6 model is currently in the process of being submitted for publication. It is also worth mentioning that to our knowledge, this is the very first time that a modelling approach is applied to the study of the palmitoylation networks. Moreover, the sub network of calnexin palmitoylation cascade is the first step towards the reconstruction of the entire cellular palmitoylation network.

Chapter 2 - State of the art

In the cellular environment, proteins are subjected to a wide variety of modifications that can occur during the lifetime of a peptide. In particular, there is a class of modifications that occur on the amino acids side chains of proteins that have been shown to be fundamental to the regulation of many protein properties like stability, localization, function and activation. This class of modifications are called post translational modifications (PTMs) [1,2]. Among PTMs there are different modifications that are reversible, these are of particular interest because they allow cells to regulate processes in time and in space [25–29]. Some of these reversible modifications have already been extensively characterized, like for example phosphorylation, a PTM with a very high number of substrates and a huge impact on many fundamental cellular functions, like metabolism and intra/extra cellular communication [3,30,31]. Not all the modifications have been equally studied, and the function and regulation of many different PTMs is still obscure. Amongst these, S-palmitoylation is gaining more and more attention due its emerging role in signalling.

2.1 What is palmitoylation?

S-palmitoylation consists in the attachment of a palmitate molecule to the SH-group of a cysteine residue via a thioester bond [32]. The mechanism of the enzymatic reaction of palmitoylation is shown in Figure 2.1.

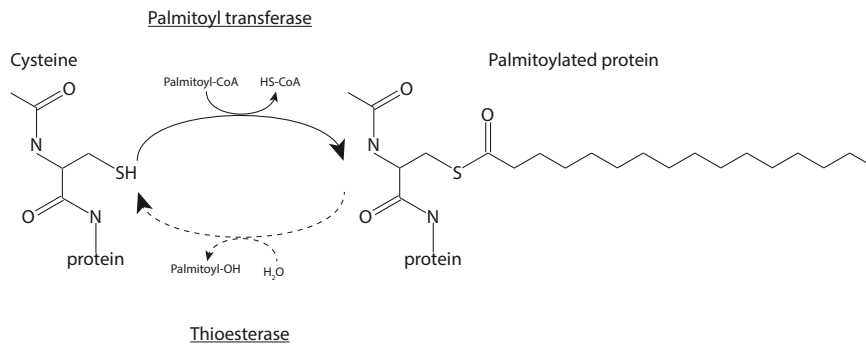


Figure 2.1. Reaction mechanism of palmitoylation. A palmitate molecule is attached to the SH group of a cytosolic cysteine of a protein via thioester bond. The reaction of palmitate attachment is catalysed by a class of proteins called Palmitoyltransferases (DHHs). The opposite reaction of palmitate removal is catalysed by another type of enzymes called Acyl protein thioesterases (APTs). Interestingly different DHHs and APTs need to be palmitoylated themselves in order to be active.

This modification has a unique property among the lipid modifications, it is in fact the only one that is reversible [33]. The attachment of the palmitate is catalysed by a class of enzymes called palmitoyltransferases (DHHs), while the opposite reaction of depalmitoylation is performed by a family of enzymes named Acyl protein thioesterases. Different studies have addressed the role of palmitoylation in regulation of cellular functions, especially regarding protein folding, localization, membrane interactions, protein-protein association and enzymes regulation [15,17,20,22,29,32,34]. Although S-palmitoylation was identified more than 40 years ago [35,36], our understanding of this modification, its dynamics, its regulation and its consequences, is still rudimentary. The principal cause in the lack of knowledge comes from the technical difficulties arising during the analysis of palmitoylated proteins and the enzymes that catalyse this PTM. In fact, since these enzymes are membrane bound, it is difficult to isolate and purify them. Moreover, the number of protocols available to analyse palmitoylation dynamics have been very limited; only recently new techniques allowed to improve our capacity of observing palmitoylation dynamics *in vivo*. The most common techniques for analysis of palmitoylated proteins include the use of radiolabelled amino acids or palmitate following pulse-chase protocols [20]. Labelled amino acids allow to monitor the dynamics of synthesis/degradation of palmitoylation substrates, observing both the fractions of modified and unmodified proteins. Radiolabelled palmitate, instead, allow us to observe only the palmitoylated fraction of a protein, giving information on the

palmitoylation/depalmitoylation dynamics of the protein. These two techniques have been useful in the past to characterize the dynamics of N-Ras palmitoylation [37] or the modulation of G α subunit palmitoylation [38]. Radiolabelling techniques are continuing to prove themselves useful as they were recently used to analyse palmitoylation of the ER chaperon calnexin [39], and its palmitoylating enzyme DHHC6 (see Chapter 5 -). Today, different new techniques have been developed to analyse palmitoylation at the protein level like acyl-biotin exchange and click chemistry, which are more sensitive with respect to radio labelling. In fact, this new techniques allowed the characterization of the cellular palmitoylome in different species, including mammalian cells [16,18,21,40,41]. This has given us an idea of the extension and complexity of the palmitoylation network. At the time of writing 19 palmitoyl-proteome studies have been published, which results have been collected in a comprehensive database of protein palmitoylation called “Swisspalm” [16].

2.2 Palmitoylation in humans

Analysis of the Swisspalm database reveals that at least 12% of the human proteins may undergo palmitoylation, demonstrating the high impact of this PTM at the cellular level. UniProt annotations allow to better characterize the pool of palmitoylated protein. 45% are membrane proteins, while cytosolic proteins constitute only 30% of the pool. Palmitoylation seems to affect especially those organelles that part of the endomembrane systems like the endoplasmic reticulum, Golgi apparatus, lysosomes and endosomes [16]. Interestingly the remaining 25% of the palmitoylated proteins are localized within the nucleus, raising the interesting possibility that palmitoylation promote reversible interaction with the nuclear membrane and/or proteins [16].

2.3 The role of palmitoylation in the regulation of cellular functions

The role of palmitoylation is still not completely understood, but different studies observed its effects on diverse cellular functions, like signalling capacity [21–23], protein localization and trafficking [34,42,43], modulation of enzymes activity [24], membranes interactions [44], and regulation of protein stability [39,45,46]. The wide variety of functions affected by palmitoylation demonstrate its role as a regulatory modification. Until now five main consequences of palmitoylation have been identified, but is still unclear how these correlate with the changes observed at the system level [15]. The most straightforward effect of palmitate attachment is that it confers hydrophobicity to proteins [15,32,47], allowing reversible interactions with the membranes or with membrane proteins. This property allows soluble proteins, which are normally hydrophilic, to increase their hydrophobicity promoting their ability of interacting temporary with the membranes. Interactions of soluble proteins with the membrane imply a multistep process, since the enzymes that catalyse palmitoylation are membrane bound and therefore palmitoylation cannot occurs directly on cytosolic proteins. Before palmitoylation occurs, generally soluble proteins undergo another lipid modification like N-myristoylation or prenylation. This first modification allows soluble proteins to gain some degree of hydrophobicity prior to palmitoylation, allowing proteins to get nearer to the membrane and start a transient, weak membrane-protein interaction. Once a protein is close to the membrane, it gains access DHHCs enzymes which will palmitoylate the protein, stabilizing its interaction with the membrane until the palmitate is removed [15,20]. Palmitoylation have been proven to be essential for a stable membrane association of many soluble proteins like Ras and G α subunit [48–50]. The increase in hydrophobicity and the consequent membrane association is something that affects most soluble proteins, while this may look as a minor effect on the fate of a protein, governing the localization of a protein in the cell can have drastic consequences at the network

level. For example, one of the fundamental requirements for enzymatic reactions to happen is co-localization of the enzyme with its substrate. In these sense, palmitoylation can promote or suppress the activity of an enzyme by enriching or sequestering the enzyme/substrate at a particular membrane.

Another direct consequence of palmitoylation that often affects membrane proteins are changes in the conformation of proteins structure. Palmitoylation of membrane protein often occurs on cysteines adjacent to the transmembrane domain [15]. A study by Joseph and Nagaraj [51] has shown that palmitoylation of cysteines near the transmembrane domain would cause the embedding of those residue deep in the membrane, causing a twist of the transmembrane domain (TMD) of the protein that would change its orientation and length. This changes in conformation can have effects on protein localization as shown in [42]. In this work, the transmembrane domain (TMD) of LRP6, a key component of the LRP5/LRP6/Frizzled co-receptor, was shortened (data not published). What was observed is that shortening relieved ER retention in the absence of palmitoylation, while in the presence of palmitoylation shortening led to ER retention, which is consistent with an effect of palmitoylation on the TMD of LRP6. Among the mechanistic consequences of palmitoylation we also find the ability of association with specific membrane domains like lipid-rafts, cholesterol and sphingolipid rich membrane domains [15]. It is indeed proved that palmitoylation of many proteins promoted their association with lipid rafts [52]. Palmitoylation is also responsible for protein-protein interactions; this is especially true for membrane proteins. A recent example is the one of the ER chaperone calnexin. Its palmitoylation is required for the formation of a super complex “translocon-ribosome-calnexin” [34]. The formation of this complex allows newly synthesized proteins to be immediately processed by enzymes and chaperones that helps the protein reach the final conformation [34]. Interestingly, It has been observed that palmitoylation can also have the opposite effect of interfering with protein interaction [15]. For example in the case of the neural cell adhesion molecule 140 (NCAM140) palmitoylation negatively regulates the hemophilic dimerization of extracellular NCAM domains [34].

Surprisingly palmitoylation seems also to interact with different PTMs, in particular with ubiquitination. It was found that palmitoylation deficient LRP6 undergoes ubiquitination near its palmitoylation sites [42]. Another similar example is constituted by TEM8, a protein which function is not clearly understood yet, but it is known to be the receptor for the bacillus anthracis toxins, the causative agent of anthrax disease. The palmitoylated deficient version of this protein has been found to be ubiquitinated near the palmitoylation sites, causing premature degradation [46]. Since different palmitoylation deficient proteins have been found to be ubiquitinated and prematurely degraded, palmitoylation is generally associated with a stabilization effect. Interplay between palmitoylation and phosphorylation has also been observed. A clear case of interaction have been observed for calnexin; a mutant version of the protein which cannot be phosphorylated led to an increase of 200% in the level of palmitoylation, and a consistent increase have also been observed in the rate of palmitate incorporation [39]. This observation suggests that palmitoylation of calnexin is under the negative control of serine phosphorylation.

2.4 Palmitoylation cycle: DHHCs and APTs

As stated earlier the reaction of S-palmitoylation is catalysed by a family of enzymes called palmitoyltransferases. So far 23 different palmitoyltransferases have been identified in mammals. This class of enzymes is characterized by the presence of a conserved DHHC (aspartate-histidine-histidine-cysteine) motif within a cysteine rich domain which is the catalytic site of these enzymes [53] and is directly implicated in the attachment of the palmitate. Mutation in the DHHC sequence abolish the ability of DHHCs to catalyse palmitoylation [39,54,55]. DHHC proteins are transmembrane proteins that can span through the membrane at least four times, with the DHHC domain normally localized between the second and the third transmembrane domain. A schematic representation of the fundamental structure of a DHHC enzyme

is visible in Figure 2.2. These enzymes are mainly localized on the ER membrane and the Golgi, but some are also found on the plasma membrane [56].

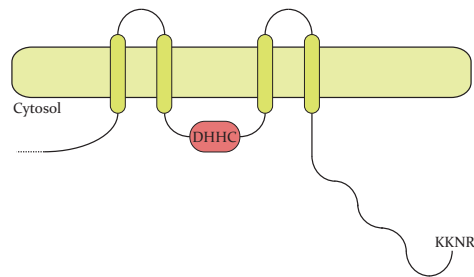


Figure 2.2. General structure of a DHC enzyme. DHC enzymes are transmembrane protein that span through the membrane at least for times. The catalytic site is a DHC motif usually located in between the 2nd and 3rd transmembrane domain, on the cytoplasmic side. On the C-terminal domain we find a localization signal, in this representation the KKNR signal is targeting the DHC to the endoplasmic reticulum. Depending on the DHC we can find different additional domain usually localized between the 4th transmembrane domain and the end of the cytoplasmic tail. Like for example the SH3 like domain in DHC6.

It is actually not clear why there are so many DHC proteins and what their substrate specificity is, but their large number and precise localization to different membrane compartments suggest that palmitoylation is a tightly regulated network with hundreds of substrates taking part in the most diverse cellular functions [15,20]. At the present time, a consensus sequence characterizing palmitoylation sites has not been found even if some prediction algorithms show increasingly successful results [16,57]. These algorithms are based on the recognition of some common features that belongs to palmitoylable cysteines, for example modifiable residues tends to localize near to hydrophobic parts of proteins. In soluble proteins, palmitoylable cysteine are often found next to myristoylation or prenylation sites, since these lipid modifications help proteins to get nearer to the membrane and to interact with DHCs enzymes. In transmembrane proteins, it is common to find palmitoylation sites just next to the transmembrane domain [20], like in the ER chaperone calnexin [58] or the anthrax toxin receptor CMG2 [46].

In regard to the specificity towards substrate we can divide DHCs in two classes, in fact, while some palmitoyltransferases are substrate-specific, others modify a broad range of substrate [59,60]. Intriguingly, DHC proteins can have overlapping

substrate specificities, meaning that the palmitoylation cycle of some substrates can be regulated by more than one DHHC enzyme. In the case of SNAP25, a protein with a key role in neuronal exocytosis pathways, the membrane association of this protein is controlled by DHHC3, 7, 17 and 15 [61]. Another good example is for one of the $G\alpha$ subunits, similarly to SNAP25, the membrane association of this protein is controlled by both DHHC3 and 7 [62].

The regulation of the palmitoyltransferase expression and activity, is still not well understood at this time. A review of the knowledge about each single DHHC can be found in [63]. Interestingly a large-scale proteomic analysis of palmitoylated proteins revealed that different DHHCs, (5,6 and 8) are themselves palmitoylated [41], suggesting that the activity of a subset of DHHCs enzymes may be regulated by palmitoylation itself, adding another level of complexity to the regulation mechanisms of this modification. Other form of regulations may require formation of complex with other protein, like it has been shown in palmitoylation of Ras in yeast and mammals [54,64]. In that study the authors show that Ras palmitoylation require the interaction between DHHC9 and GCPI6.

Unfortunately, if our knowledge on DHHCs is poor, what we know on depalmitoylation enzymes is even less. The enzymes responsible for removal of palmitate belong to a class of proteins named thioesterases [65]. Two of these enzymes, protein palmitoyl thioesterase 1 (PPT1) and PPT2 are found in lysosomes and depalmitoylate proteins that are being degraded [66]. Two other enzymes called acyl protein thioesterase 1 (APT1) and APT2 are localized in the cytoplasm and seems to have a more active role in the regulation of the palmitoylation cycle of a variety of substrates. APT1 have been show to take part to the regulation of the palmitoylation cycle of proteins like $G\alpha$ subunits and H/N-Ras [66,67]. HEK293 cell show an increased rate of removal of radiolabelled palmitate from Galpha protein after overexpression of APT1 in pulse-chase experiments [67]. Analysis of an N-ras protein in which the palmitoyl group was attached by a noncleavable thioether linkage revealed that the cycle of palmitoylation/depalmitoylation of N-Ras and the correct

localization of its palmitoylated form is important to determine the correct cellular distributions [66,67]. Conversely, APT2, has been identified as the thioesterase of the growth associated protein GAP43 and H-Ras [68]. Quite recently, a large profile study using broad spectrum serine lipase have highlighted that different members of the serine hydrolase family are capable of depalmitoylation. This may suggest that we still haven't identified all the enzymes responsible for protein depalmitoylation, and therefore a consistent part of the dynamic network of protein palmitoylation still need to be uncovered.

2.5 Study palmitoylation with the help of a modelling approach

This thesis is focused on the study of palmitoylation using a Systems biology approach. Based on experiments we developed mathematical models of the palmitoylation process of different modified proteins, to understand the biological behaviour of this PTM at the system level.

Kinetic modelling has been extensively used to improve the comprehension of complex cellular networks. Particularly successful results have been obtained in the study of signalling networks. These networks are normally characterized by a moderate number of proteins, and their topology may appear simple at a first look, but the large number of biological responses that these network can trigger, and the pleiotropic effect on a number of cellular function reveal the hidden complexity of these models, often characterized by complex non-linear interaction and the existence of proteins in multiple states thanks to the presence of post translational modification sites [69] Recently the combined approach of experiments and modelling is revealing a successful strategy to characterize the dynamics of signalling network, especially regarding the study of post translational modifications.

2.6 Study palmitoylation with the help of modelling

Kinetic modelling has been extensively used to improve the comprehension of cellular networks. Particularly successful results have been obtained in the study and comprehension of signalling networks, which are normally characterized by a moderate number of proteins, but high complexity. The topology of these systems may appear simple at a first look, but the large number of biological responses that these networks can trigger, along with pleiotropic effects on a number of cellular functions reveal the hidden complexity of these networks [69]. For example in PC12 cells sustained activation of ERK triggers neuronal differentiation, while a transient activation is required for proliferation [70]. Even if we know the possible outcomes of the network, we still don't fully understand how these correlates with the dynamics of the signalling network [69]. In this regard, systems biology is used to study how the intrinsic network dynamics allow to trigger different biological responses through the use of mathematical formalism for the description of those systems.

Mathematical models allow the study of protein networks on different levels. A common approach to model analysis consist in perturbing its components and predict the possible response of the network. Perturbations can be generated at different levels and for each protein of the model, allowing the individuation of the principal components, meaning those parts of the network which small perturbations have a big impact on the dynamics of the system [69]. Another remarkable aspect of mathematical modelling is that it can reveal emergent properties, which are features of the network that rise from its functioning as a whole, and they are not observed on single proteins [69]. A famous example is the ability of certain proteins, participating in signalling networks, to be activated in a switch-like manner. This property is known as ultrasensitivity and its function is to filter background noise and avoid accidental activation of signalling cascades. A pathway showing this property is the ERK activation system in *Xenopus oocytes* [71]. In this work *Xenopus oocytes* where threated with progesterone to study the ERK activation pathway and subsequent

maturation of the oocytes. Interestingly, while the population of oocytes showed a linear activation of ERK in response to increasing levels of progesterone, individual eggs exhibited an all or nothing response [69], [71]. The linear response can be simply explained by an increase in the frequency of activated cells. Instead, the non or all response of single cells rise instead from a more interesting mechanism of ERK activation. In this pathway after activation of the growth factor (GF) receptors the kinase MEK, which have two phosphorylation sites, gets modified on both sites and subsequently phosphorylates ERK twice, triggering DNA transcription [71]. When MEK phosphorylates ERK it modifies only one site at the time, therefore, two distinct enzymatic reactions of phosphorylation are necessary to fully phosphorylate ERK. This imply that a poor stimulus of the pathway will generate a pool of mono phosphorylated ERK, which won't be active. The triggering of this pathway require that the initial stimulus have a certain level of strength in order to activate all the proteins in the pathway and generate a response. As consequence the ERK activation adopt a sigmoidal shape and its activation become switch-like [69][71]. In such a complex network of interactions intuition is not enough to explain the difference of behaviour observed between the population and the single cell response. Moreover, ultrasensitivity is not a property that can be studied by simply analyse the components of the network, since it can only be observed when we consider the network as a whole. Therefore, only a mathematical approach will be able to predict exactly how the network is activated upon a stimulus of the GF receptors [69]. The examples above show how the combined approach of experiments and modelling is revealing a successful strategy to characterize the dynamics of signalling networks, and to explain unexpected behaviours that are encoded in the design of biological networks [69].

This work is focused on the study of palmitoylation using a systems biology approach. Based on experiments we plan to develop mathematical models of different palmitoylated proteins belonging to the palmitoylation cascade of the ER chaperone calnexin, with the aim of studying their dynamics of palmitoylation. First we plan to design a single protein model describing the double modification process of calnexin.

This model will be first used to validate the reaction mechanisms that we choose to describe the enzymatic reaction of palmitoylation. If the proposed mechanism will manage to correctly describe the dynamics of calnexin palmitoylation, the model will be investigated to uncover the consequences of calnexin modification. We will focus on the study of the dynamics of modification of the two calnexin sites: Are the sites modified sequentially? Is palmitoylation dynamic for calnexin? Which percentage of population is palmitoylated in steady state and how long does it take to palmitoylate calnexin? Moreover, we will investigate how palmitoylation affects the stability of the protein.

A similar approach will be applied also to DHHC6, the palmitoyltransferase responsible for palmitoylation of calnexin. This enzyme has three palmitoylation sites, and experiments have shown that DHHC6 palmitoylation is fundamental for enzyme activity, stability and localization. A single protein model will contribute to investigate more in detail the effects of palmitoylation on the properties of this enzyme. In particular, we are interested in understanding if the modification of each site have the same consequences on the enzyme. We will also monitor the dynamics of incorporation of palmitate in DHHC6, determining if the three sites are equally palmitoylated. Studies on the stability of the different palmitoylation patterns of DHHC6 will also reveal how the three sites contribute to determine the stability of the enzyme. We will observe if palmitoylation of DHHC6 is dynamic, an important feature for temporary activation and deactivation of the enzyme. Prediction on steady state distribution will show how much enzyme is active in steady state or if there are preferentially palmitoylated states. All these findings will be useful to increase our understanding of how palmitoylation regulates the activity and stability of DHHC6.

In the final part of the project the two models of calnexin and DHHC6 will be used as starting point for the reconstruction of the palmitoylation cascade of calnexin. In this model we will include all the principal proteins with a role in calnexin palmitoylation. This will allow to study the dynamics of palmitoylation at the system level. We will first investigate how palmitoylation of DHHC6 affects palmitoylation of calnexin,

trying to understand how the different DDHC6 sites contribute to modulate its activity. We will determine which parameters of the model are fundamental in determining the dynamics of palmitoylation observed in steady states, revealing which part of the network are more important for its correct functioning. Perturbing the model components will help to determine their function, possibly revealing emergent properties of the cascade. Another interesting point would be to determine what happens when we increase or decrease the activity of those enzymes that regulate the cycles of palmitoylation/depalmitoylation of the proteins belonging to the palmitoylation cascade of calnexin.

All together the steps proposed in this project will greatly contribute in the understanding of palmitoylation, helping to unravel the complexity of palmitoylation dynamics both at the protein and system level.

Chapter 3 - Material and Methods

In this section we focus first on the tools that we developed to set up and analyse the different palmitoylation models. We start illustrating how we defined the network topology for a palmitoylation reaction [72], how this was translated into a mathematical model, and the approximation that we applied to reduce model complexity.

Next we talk about the tools created in order analyze the models. Part of this section is dedicated to the *in-silico* labelling technique that was developed *ad-hoc* to be able to estimate parameters from radio labelling *in-vivo* experiments. In this section we also talk about genetic algorithm, which, in combination with *in-silico* labelling, was used to perform parameter estimation.

In the model analysis section, we briefly introduce the concept of local and global sensitivity and how we use them to analyse the models. Others tools, like single molecule tracking with stochastic simulations are also part of this section. The Chapter ends with an introduction on Rule-based modelling, and how it is used to reconstruct a small network of palmitoylation starting from the single protein models developed previously.

3.1 Modeling of palmitoylation networks: Define the network topology

The core of all our models is based on the well-established model of protein phosphorylation illustrated in [72]. The set of reactions described in Goldbeter paper is used to describe a single protein palmitoylation event, as shown in Figure 3.1.

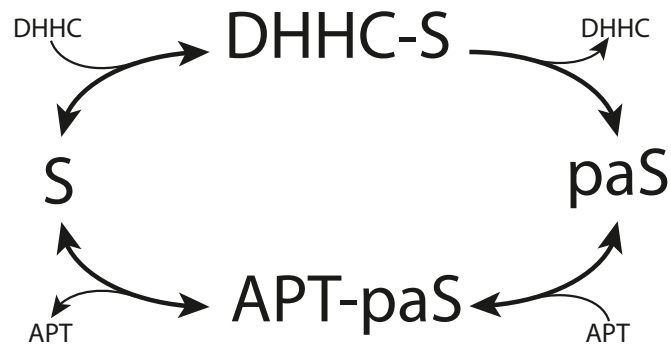


Figure 3.1 Generic protein palmitoylation/depalmitoylation cycle. In a generic reaction of palmitoylation a substrate S binds to a palmitoyltransferases (DHHC) enzyme forming the complex DHHC-S. DHHC catalyse the attachment of a palmitate molecule pa on the substrate S and release the palmitoylated substrate paS . In the opposite reaction of depalmitoylation paS bind to an Acyl protein thioesterase (APT) forming the complex APT- paS . The palmitate molecule is then removed and the substrate released.

Multiple modification events are modelled replicating this subunit for each palmitoylation site present on the protein. An example showing the cycle of palmitoylation/depalmitoylation of calnexin is shown in Figure 3.2.

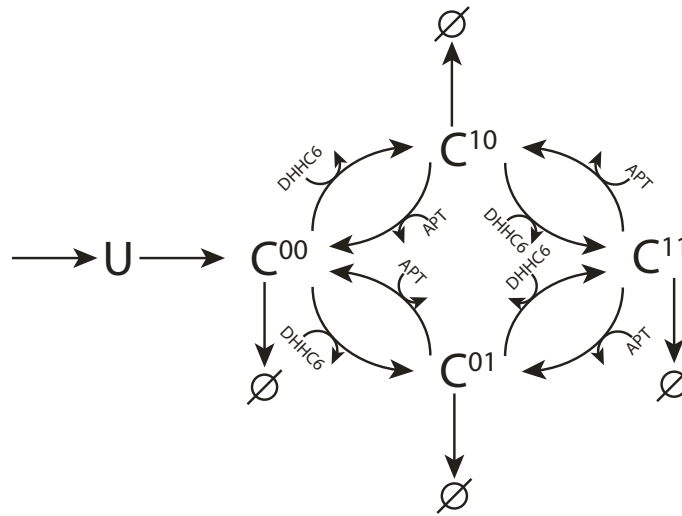


Figure 3.2 Model of calnexin palmitoylation/depalmitoylation cycle. We assume that calnexin is first synthesized by the ribosome from its mRNA to the corresponding peptide and inserted into the membrane (U). Unfolded calnexin goes through a process of folding, forming folded calnexin, which then shows the two sites available for palmitoylation (C^{00}). These two sites can be palmitoylated by DHHc6. The first palmitoylation can occur on both sites; C^{10} and C^{01} denote the two different palmitoylation states. After the first PTM calnexin can undergo another palmitoylation event still catalysed by DHHc6; C^{11} represents the double palmitoylated calnexin. Since palmitoylation is reversible the palmitate can be removed from the two sites. The removal of palmitate requires a different type of enzyme, namely Acyl Protein Thioesterases (APTs).

Different experiments on mutant proteins substrates of palmitoylation in which one or more palmitoylation sites were removed have shown that palmitoylation of a particular site is independent from the others [34] (see also Chapter 5 -). Because of this property, during the definition of the network we decided to include all the possible palmitoylation state of a protein, and all the connections between these states. For example in the model shown in Figure 3.2, calnexin can be palmitoylated on two sites independently. This means that calnexin can exist in four different palmitoylation states (C^{00} , C^{10} , C^{01} and C^{11}). Now, since the modification of these sites is independent, the transition from C^{00} to C^{11} can happen in two different ways ($C^{00} \rightarrow C^{10} \rightarrow C^{11}$ or $C^{00} \rightarrow C^{01} \rightarrow C^{11}$). Therefore, both the options have been included in the model, and the same is done for all the models developed.

In the Goldbeter work, the model was mathematically described using mass action terms. In this work, since we are dealing with multiple modification events, and because the final aim of the project is to provide an extended representation of the palmitoylation network, a mass action approach would be prohibitive in terms of computation, number of parameters and time. Therefore, we choose to describe each

enzymatic step using the “quasi-steady state approximation” (QSSA). Although this assumption can be valid for many enzymatic reactions *in vitro*, it fails to describe those reactions in which the protein and the enzyme have comparable concentrations, as often happen inside the cell. In fact the validity of QSSA is guaranteed only until, i.e. when the enzyme concentration is low with respect to the substrate $\frac{E_T}{K_M + S_T} \ll 1$ [2, 3].

Since it is known that in the palmitoylation network different enzymes and substrates have comparable concentrations [21,40,41,73,74], we solved the problem adopting what is called “total quasi-steady state approximation” (tQSSA), well described in [75] and [76]. This approximation has been successfully used in models describing multiple phosphorylation modification in signalling cascades [76]. Moreover, when we have to deal with cellular processes, Tzafiriri et al. [77] have shown that the tQSSA approximation is valid also when the enzyme substrate concentrations are comparable [78].

3.2 Application of tQSSA to the enzymatic reactions of palmitoylation

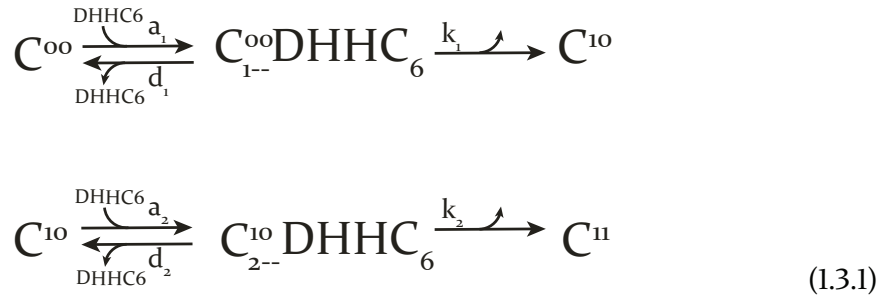
Describing a signalling model with mass action kinetics can lead to very complicated equations, depending on the number of species and the interactions among them. Moreover, the kinetic rate constants used in mass action kinetics, like binding rates or catalytic constants, are difficult or impossible to measure experimentally. When dealing with enzymatic reactions, as typically observed in signalling models, we can reduce the complexity of the model by applying an approximation to the mass action kinetics called quasi steady state approximation (QSSA). QSSA allows to simplify the equations describing enzymatic reactions and to group mass action kinetic rate constants into more meaningful parameters that can even be measured

experimentally. The resulting equations have the form of Michaelis-Menten kinetics. This approximation is based on the assumption that in an enzymatic reaction the enzyme-substrate complex at the beginning of the reaction immediately reaches a chemical equilibrium after the reaction start.

In this section we show the step-by-step application of tQSSA on the mass action model of calnexin palmitoylation visible in Figure 3.2. The exact same procedure is applied to each model developed in this project.

To facilitate the comprehension, we first demonstrate how tQSSA is applied to a double step of palmitoylation, without taking into account depalmitoylation and degradation reactions. This system is identical to the competitive system described in [75], therefore the exact same procedure can be used to apply tQSSA to the calnexin model.

Specifically, consider the following successive reactions of calnexin modification that lead to double palmitoylation of the protein starting from the non palmitoylated form:



Where C^{00} represent folded but not palmitoylated calnexin, C^{10} and C^{11} denote the single and double palmitoylated states respectively. $C^{00}\text{DHHC}_6$ denotes the complex of non-palmitoylated calnexin with the palmitoyltransferase 6, which is responsible for the modification of both sites. $C^{10}\text{DHHC}_6$ represent a similar complex in which calnexin is already palmitoylated on site 1. This reaction scheme resembles closely the system presented in equation 14 of [75]. In both cases we have two reactions catalysed by the same enzyme, therefore the two substrates are competing to be modified. The only difference in the two reaction schemes is that in our case the product of the first

reaction is substrate in the second reaction, but we show that this does not make any difference in the application of tQSSA.

The reaction scheme in (1.3.1) is governed by the coupled ODE (1.3.2)-(1.3.6) below:

$$\frac{dC^{oo}}{dt} = -a_1 \cdot DHHC_6 \cdot C^{oo} + d_1 \cdot C_{--}^{oo} DHHC_6, \quad (1.3.2)$$

$$\frac{dC_{--}^{oo} DHHC_6}{dt} = a_1 \cdot (DHHC_6 \cdot C^{oo} - K_1^M \cdot dC_{--}^{oo} DHHC_6), \quad K_1^M = \frac{d_1 + k_1}{a_1} \quad (1.3.3)$$

$$\frac{dC^{10}}{dt} = k_1 \cdot C_{--}^{oo} DHHC_6 - a_2 \cdot DHHC_6 \cdot C^{10} + d_2 \cdot C_{--}^{10} DHHC_6, \quad (1.3.4)$$

$$\frac{dC_{--}^{10} DHHC_6}{dt} = a_2 \cdot (DHHC_6 \cdot C^{10} - K_2^M \cdot dC_{--}^{10} DHHC_6), \quad K_2^M = \frac{d_2 + k_2}{a_2} \quad (1.3.5)$$

$$DHHC_6^{Tot} = DHHC_6 + C_{--}^{oo} DHHC_6 + C_{--}^{10} DHHC_6. \quad (1.3.6)$$

Now, in order to apply tQSSA we define new quantities, the total substrates:

$$\begin{aligned} \overline{C^{oo}} &= C^{oo} + C_{--}^{oo} DHHC_6 \\ \overline{C^{10}} &= C^{10} + C_{--}^{10} DHHC_6 \end{aligned} \quad (1.3.7)$$

We then rewrite equations (1.3.2)-(1.3.6) in terms of total substrates

$$\frac{d\overline{C^{oo}}}{dt} = -k_1 \cdot C_{--}^{oo} DHHC_6, \quad (1.3.8)$$

$$\frac{dC_{--}^{oo} DHHC_6}{dt} = a_1 \left((DHHC_6^{Tot} - C_{--}^{oo} DHHC_6 - C_{--}^{10} DHHC_6) \cdot (\overline{C^{oo}} - C_{--}^{oo} DHHC_6) - K_1^M \cdot C_{--}^{oo} DHHC_6 \right), \quad (1.3.9)$$

$$\frac{d\overline{C^{10}}}{dt} = k_1 \cdot C_{--}^{oo} DHHC_6 - k_2 \cdot C_{--}^{10} DHHC_6, \quad (1.3.10)$$

$$\frac{dC_{--}^{10}DHHC_6}{dt} = a_2 \left((DHHC_6^{Tot} - C_{--}^{00}DHHC_6 - C_{--}^{10}DHHC_6) \cdot (C_{--}^{10} - C_{--}^{00}DHHC_6) - K_2^M \cdot C_{--}^{10}DHHC_6 \right). \quad (1.3.11)$$

Equations (1.3.8)-(1.3.11) are identical to equations 21a-b in [75]. Then, we apply the quasi-steady-state assumption

$$\begin{aligned} \frac{dC_{--}^{00}DHHC_6}{dt} &\approx 0, \\ \frac{dC_{--}^{10}DHHC_6}{dt} &\approx 0. \end{aligned} \quad (1.3.12)$$

Which gives

$$C_{--}^{00}DHHC_6 = DHHC_6^{Tot} - C_{--}^{10}DHHC_6 \left(1 + \frac{K_2^M}{C_{--}^{10} - C_{--}^{10}DHHC_6} \right), \quad (1.3.13)$$

$$C_{--}^{10}DHHC_6 = DHHC_6^{Tot} - C_{--}^{00}DHHC_6 \left(1 + \frac{K_1^M}{C_{--}^{00} - C_{--}^{00}DHHC_6} \right). \quad (1.3.14)$$

equations (1.3.13) and (1.3.14) are equivalent to 22a and 22b in [75].

Substituting (1.3.14) into (1.3.13) leads to the following equation for complex $C_{--}^{00}DHHC_6$ and then complex $C_{--}^{10}DHHC_6$ can be calculated from (1.3.14).

$$C_{--}^{00}DHHC_6 = DHHC_6^{Tot} - \left(DHHC_6^{Tot} - C_{--}^{00}DHHC_6 \cdot \left(1 + \frac{K_1^M}{C_{--}^{00} - C_{--}^{00}DHHC_6} \right) \right) \cdot \left(1 + \frac{K_2^M}{C_{--}^{10} - \left(DHHC_6^{Tot} - C_{--}^{00}DHHC_6 \cdot \left(1 + \frac{K_1^M}{C_{--}^{00} - C_{--}^{00}DHHC_6} \right) \right)} \right) \quad (1.3.15)$$

It is easy to notice that equation (1.3.15) is identical to equation (23) from [75], therefore solving (1.3.15) is equivalent to finding the roots of the third degree polynomial (24) in [75]. The polynomial is solved following the exact same procedure as in [75], (equations (24-26)), finally obtaining:

$$\begin{aligned}
 \frac{d\overline{C}^{oo}}{dt} &= -\frac{V_f \cdot \overline{C}^{oo}}{K_1^M \cdot \left(1 + \frac{\overline{C}^{10}}{K_2^M}\right) + \overline{C}^{oo} + DHHC_6^{Tot}}, \\
 \frac{d\overline{C}^{10}}{dt} &= \frac{V_f \cdot \overline{C}^{oo}}{K_1^M \cdot \left(1 + \frac{\overline{C}^{10}}{K_2^M}\right) + \overline{C}^{oo} + DHHC_6^{Tot}} - \frac{V_f \cdot \overline{C}^{10}}{K_2^M \cdot \left(1 + \frac{\overline{C}^{oo}}{K_1^M}\right) + \overline{C}^{10} + DHHC_6^{Tot}}.
 \end{aligned} \tag{1.3.16}$$

The system of equations (1.3.16) contains in the denominators the terms of competing substrates for the active sites of the enzyme DHHC6. In case of multiple competing species we would have additional terms in the denominator as described by the generalized equation (27) in [75].

Equivalently, we derive the reaction rates for the reverse steps of depalmitoylation, catalyzed by a different enzyme, namely acyl protein thioesterases (APT).

Based on the above, we show the equivalent steps of applying the tQSSA in the model of calnexin palmitoylation (Figure 3.1). Figure 3.3 describes the network of calnexin's palmitoylation/depalmitoylation with elementary reaction steps that include the formation of the intermediary enzyme complexes.

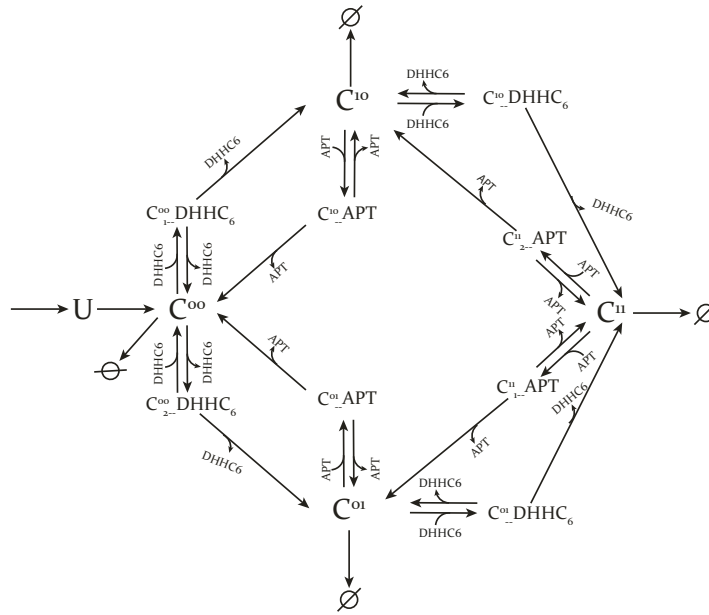


Figure 3.3. Mass action model of calnexin palmitoylation. In this model calnexin is first synthesized in the unfolded state U . The folding process brings the protein in the fully folded and not palmitoylated state C^{00} . Calnexin can then be palmitoylated on one of the two sites, ending up being modified once, either on site 1 (C^{10}) or site 2 (C^{01}). Calnexin can undergo a second palmitoylation ending up in the fully palmitoylated state C^{11} . We assume that after folding calnexin can be degraded in each state. The reaction of palmitoylation are catalysed by DHHC6, while the opposite reaction of palmitate removal by APT1.

The system of ODEs in (1.3.17) fully describes the model visible in Figure 3.3:

$$\begin{aligned}
 \frac{dU}{dt} &= k_s - k_f \cdot U \\
 \frac{dC_{-}^{00}}{dt} &= k_f \cdot U + k_{-1f}^{C_{-}^{00,1,1}} \cdot C_{-}^{00} \cdot DHHC_6 + k_{-1f}^{C_{-}^{00,2,1}} \cdot C_{-}^{00} \cdot DHHC_6 + k_{2b} \cdot C_{-}^{10} \cdot APT + k_{2b} \cdot C_{-}^{01} \cdot APT - k_{d1} \cdot C_{-}^{00} - k_{1f}^{C_{-}^{00,1,1}} \cdot C_{-}^{00} \cdot DHHC_6 - k_{1f}^{C_{-}^{00,2,1}} \cdot C_{-}^{00} \cdot DHHC_6 \\
 \frac{dC_{-}^{10} \cdot DHHC_6}{dt} &= k_{1f}^{C_{-}^{00,1,1}} \cdot C_{-}^{00} \cdot DHHC_6 - k_{-1f}^{C_{-}^{00,1,1}} \cdot C_{-}^{10} \cdot DHHC_6 - k_{2f} \cdot C_{-}^{10} \cdot DHHC_6 \\
 \frac{dC_{-}^{01} \cdot DHHC_6}{dt} &= k_{1f}^{C_{-}^{00,2,1}} \cdot C_{-}^{00} \cdot DHHC_6 - k_{-1f}^{C_{-}^{00,2,1}} \cdot C_{-}^{01} \cdot DHHC_6 - k_{2f} \cdot C_{-}^{01} \cdot DHHC_6 \\
 \frac{dC_{-}^{10}}{dt} &= k_{2f} \cdot C_{-}^{00} \cdot DHHC_6 + k_{-1b}^{C_{-}^{00,1,1}} \cdot C_{-}^{10} \cdot APT + k_{-1f}^{C_{-}^{00,1,2}} \cdot C_{-}^{10} \cdot DHHC_6 + k_{2b} \cdot C_{-}^{11} \cdot APT - k_{d2} \cdot C_{-}^{10} - k_{1b}^{C_{-}^{00,1,1}} \cdot C_{-}^{10} \cdot APT - k_{1f}^{C_{-}^{00,1,2}} \cdot C_{-}^{10} \cdot DHHC_6 \\
 \frac{dC_{-}^{01}}{dt} &= k_{2f} \cdot C_{-}^{00} \cdot DHHC_6 + k_{-1b}^{C_{-}^{00,2,1}} \cdot C_{-}^{01} \cdot APT + k_{-1f}^{C_{-}^{00,2,2}} \cdot C_{-}^{01} \cdot DHHC_6 + k_{2b} \cdot C_{-}^{11} \cdot APT - k_{d2} \cdot C_{-}^{01} - k_{1b}^{C_{-}^{00,2,1}} \cdot C_{-}^{01} \cdot APT - k_{1f}^{C_{-}^{00,2,2}} \cdot C_{-}^{01} \cdot DHHC_6 \\
 \frac{dC_{-}^{10} \cdot APT}{dt} &= k_{1b}^{C_{-}^{00,1,1}} \cdot C_{-}^{10} \cdot APT - k_{-1b}^{C_{-}^{00,1,1}} \cdot C_{-}^{10} \cdot APT - k_{2b} \cdot C_{-}^{10} \cdot APT \\
 \frac{dC_{-}^{01} \cdot APT}{dt} &= k_{1b}^{C_{-}^{00,2,1}} \cdot C_{-}^{01} \cdot APT - k_{-1b}^{C_{-}^{00,2,1}} \cdot C_{-}^{01} \cdot APT - k_{2b} \cdot C_{-}^{01} \cdot APT \\
 \frac{dC_{-}^{10} \cdot DHHC_6}{dt} &= k_{1f}^{C_{-}^{00,1,2}} \cdot C_{-}^{10} \cdot DHHC_6 - k_{-1f}^{C_{-}^{00,1,2}} \cdot C_{-}^{10} \cdot DHHC_6 - k_{2f} \cdot C_{-}^{10} \cdot DHHC_6 \\
 \frac{dC_{-}^{01} \cdot DHHC_6}{dt} &= k_{1f}^{C_{-}^{00,2,2}} \cdot C_{-}^{01} \cdot DHHC_6 - k_{-1f}^{C_{-}^{00,2,2}} \cdot C_{-}^{01} \cdot DHHC_6 - k_{2f} \cdot C_{-}^{01} \cdot DHHC_6 \\
 \frac{dC_{-}^{11}}{dt} &= k_{2f} \cdot C_{-}^{10} \cdot DHHC_6 + k_{2f} \cdot C_{-}^{01} \cdot DHHC_6 + k_{-1b}^{C_{-}^{00,1,2}} \cdot C_{-}^{11} \cdot APT + k_{-1b}^{C_{-}^{00,2,2}} \cdot C_{-}^{11} \cdot APT - k_{d3} \cdot C_{-}^{11} - k_{1b}^{C_{-}^{00,1,2}} \cdot C_{-}^{11} \cdot APT - k_{1b}^{C_{-}^{00,2,2}} \cdot C_{-}^{11} \cdot APT \\
 \frac{dC_{-}^{11} \cdot APT}{dt} &= k_{1b}^{C_{-}^{00,1,2}} \cdot C_{-}^{11} \cdot APT - k_{-1b}^{C_{-}^{00,1,2}} \cdot C_{-}^{11} \cdot APT - k_{2b} \cdot C_{-}^{11} \cdot APT \\
 \frac{dC_{-}^{01} \cdot APT}{dt} &= k_{1b}^{C_{-}^{00,2,2}} \cdot C_{-}^{01} \cdot APT - k_{-1b}^{C_{-}^{00,2,2}} \cdot C_{-}^{01} \cdot APT - k_{2b} \cdot C_{-}^{01} \cdot APT \\
 DHHC_6^{Tot} &= DHHC_6 + C_{-}^{00} \cdot DHHC_6 + C_{-}^{10} \cdot DHHC_6 + C_{-}^{01} \cdot DHHC_6 + C_{-}^{11} \cdot DHHC_6 \\
 APT^{Tot} &= C_{-}^{10} \cdot APT + C_{-}^{01} \cdot APT + C_{-}^{11} \cdot APT + C_{-}^{01} \cdot APT
 \end{aligned}$$

(1.3.17)

We now define the total substrates. For each state of palmitoylation, the total calnexin species is the sum of the free substrate and the complexes of the calnexin species to both enzymes (DHHC₆ and APT).

$$\begin{aligned}
 \overline{C^{00}} &= C^{00} + C_{-}^{00} \cdot DHHC_6 + C_{-}^{00} \cdot DHHC_6 \\
 \overline{C^{10}} &= C^{10} + C_{-}^{10} \cdot DHHC_6 + C_{-}^{10} \cdot APT \\
 \overline{C^{01}} &= C^{01} + C_{-}^{01} \cdot DHHC_6 + C_{-}^{01} \cdot APT \\
 \overline{C^{11}} &= C^{11} + C_{-}^{11} \cdot APT + C_{-}^{11} \cdot APT
 \end{aligned} \tag{1.3.18}$$

Now, we rewrite (1.3.17) in function of these new quantities:

$$\begin{aligned}
 \frac{dU}{dt} &= k_s - k_f \cdot U \\
 \frac{dC^{00}}{dt} &= k_f \cdot U + k_{2b} \cdot C_{1-}^{10} APT + k_{2b} \cdot C_{2-}^{00} APT - k_{d1} \cdot C^{00} - k_{2f} \cdot C_{1-}^{00} DHHC_6 - k_{2f} \cdot C_{2-}^{00} DHHC_6 \\
 \frac{dC_{1-}^{00} DHHC_6}{dt} &= k_{1f}^{C^{00,1,1}} \cdot \left(\overline{C^{00}} - C_{1-}^{00} DHHC_6 - C_{2-}^{00} DHHC_6 \right) \cdot \left(DHHC_6^{Tot} - C_{1-}^{00} DHHC_6 - C_{2-}^{00} DHHC_6 - C_{1-}^{10} DHHC_6 - C_{2-}^{00} DHHC_6 \right) - K_f^{M,1,1} \cdot C_{1-}^{00} DHHC_6 \quad K_f^{M,1,1} = \frac{k_{-1f}^{C^{00,1,1}} + k_{2f}}{k_{1f}^{C^{00,1,1}}} \\
 \frac{dC_{2-}^{00} DHHC_6}{dt} &= k_{1f}^{C^{00,2,1}} \cdot \left(\overline{C^{00}} - C_{1-}^{00} DHHC_6 - C_{2-}^{00} DHHC_6 \right) \cdot \left(DHHC_6^{Tot} - C_{1-}^{00} DHHC_6 - C_{2-}^{00} DHHC_6 - C_{1-}^{10} DHHC_6 - C_{2-}^{00} DHHC_6 \right) - K_f^{M,2,1} \cdot C_{2-}^{00} DHHC_6 \quad K_f^{M,2,1} = \frac{k_{-1f}^{C^{00,2,1}} + k_{2f}}{k_{1f}^{C^{00,2,1}}} \\
 \frac{dC^{10}}{dt} &= k_{2f} \cdot C_{1-}^{00} DHHC_6 + k_{2b} \cdot C_{2-}^{00} APT - k_{d2} \cdot C^{10} - k_{2f} \cdot C_{1-}^{10} DHHC_6 - k_{2b} \cdot C_{2-}^{10} APT \\
 \frac{dC_{1-}^{10}}{dt} &= k_{2f} \cdot C_{2-}^{00} DHHC_6 + k_{2b} \cdot C_{1-}^{00} APT - k_{d2} \cdot C_{1-}^{10} - k_{2f} \cdot C_{1-}^{10} DHHC_6 - k_{2b} \cdot C_{2-}^{10} APT \\
 \frac{dC_{1-}^{10} APT}{dt} &= k_{1b}^{C^{10,1,1}} \cdot \left(\overline{C^{10}} - C_{1-}^{10} DHHC_6 - C_{2-}^{10} APT \right) \cdot \left(APT^{Tot} - C_{1-}^{10} APT - C_{2-}^{00} APT - C_{1-}^{10} APT - C_{2-}^{10} APT \right) - K_b^{M,1,1} \cdot C_{1-}^{10} APT \quad K_b^{M,1,1} = \frac{k_{-1b}^{C^{10,1,1}} + k_{2b}}{k_{1b}^{C^{10,1,1}}} \\
 \frac{dC_{2-}^{10} APT}{dt} &= k_{1b}^{C^{10,2,1}} \cdot \left(\overline{C^{10}} - C_{1-}^{10} DHHC_6 - C_{2-}^{10} APT \right) \cdot \left(APT^{Tot} - C_{1-}^{10} APT - C_{2-}^{00} APT - C_{1-}^{10} APT - C_{2-}^{10} APT \right) - K_b^{M,2,1} \cdot C_{2-}^{10} APT \quad K_b^{M,2,1} = \frac{k_{-1b}^{C^{10,2,1}} + k_{2b}}{k_{1b}^{C^{10,2,1}}} \\
 \frac{dC_{1-}^{10} DHHC_6}{dt} &= k_{1f}^{C^{10,1,2}} \cdot \left(\overline{C^{10}} - C_{1-}^{10} DHHC_6 - C_{2-}^{10} APT \right) \cdot \left(DHHC_6^{Tot} - C_{1-}^{10} DHHC_6 - C_{2-}^{10} DHHC_6 - C_{1-}^{10} DHHC_6 - C_{2-}^{10} DHHC_6 \right) - K_f^{M,1,2} \cdot C_{1-}^{10} DHHC_6 \quad K_f^{M,1,2} = \frac{k_{-1f}^{C^{10,1,2}} + k_{2f}}{k_{1f}^{C^{10,1,2}}} \\
 \frac{dC_{2-}^{10} DHHC_6}{dt} &= k_{1f}^{C^{10,2,2}} \cdot \left(\overline{C^{10}} - C_{1-}^{10} DHHC_6 - C_{2-}^{10} APT \right) \cdot \left(DHHC_6^{Tot} - C_{1-}^{10} DHHC_6 - C_{2-}^{10} DHHC_6 - C_{1-}^{10} DHHC_6 - C_{2-}^{10} DHHC_6 \right) - K_f^{M,2,2} \cdot C_{2-}^{10} DHHC_6 \quad K_f^{M,2,2} = \frac{k_{-1f}^{C^{10,2,2}} + k_{2f}}{k_{1f}^{C^{10,2,2}}} \\
 \frac{dC^{11}}{dt} &= k_{2f} \cdot C_{1-}^{10} DHHC_6 + k_{2f} \cdot C_{2-}^{10} DHHC_6 - k_{d3} \cdot C^{11} - k_{2b} \cdot C_{1-}^{11} APT - k_{2b} \cdot C_{2-}^{11} APT \\
 \frac{dC_{1-}^{11} APT}{dt} &= k_{1b}^{C^{11,1,2}} \cdot \left(\overline{C^{11}} - C_{1-}^{11} APT - C_{2-}^{11} APT \right) \cdot \left(APT^{Tot} - C_{1-}^{11} APT - C_{2-}^{00} APT - C_{1-}^{11} APT - C_{2-}^{11} APT \right) - K_b^{M,1,2} \cdot C_{1-}^{11} APT \quad K_b^{M,1,2} = \frac{k_{-1b}^{C^{11,1,2}} + k_{2b}}{k_{1b}^{C^{11,1,2}}} \\
 \frac{dC_{2-}^{11} APT}{dt} &= k_{1b}^{C^{11,2,2}} \cdot \left(\overline{C^{11}} - C_{1-}^{11} APT - C_{2-}^{11} APT \right) \cdot \left(APT^{Tot} - C_{1-}^{11} APT - C_{2-}^{00} APT - C_{1-}^{11} APT - C_{2-}^{11} APT \right) - K_b^{M,2,2} \cdot C_{2-}^{11} APT \quad K_b^{M,2,2} = \frac{k_{-1b}^{C^{11,2,2}} + k_{2b}}{k_{1b}^{C^{11,2,2}}}
 \end{aligned}
 \tag{1.3.19}$$

For the ODE of each complex we sum the corresponding ODEs of the species appearing in (1.3.18). Next, we apply the quasi-steady-state assumption:

$$\begin{aligned}
 \frac{dC_{1-}^{00} DHHC_6}{dt} &\approx 0 \\
 \frac{dC_{2-}^{00} DHHC_6}{dt} &\approx 0 \\
 \frac{dC_{1-}^{10} APT}{dt} &\approx 0 \\
 \frac{dC_{2-}^{10} APT}{dt} &\approx 0 \\
 \frac{dC_{1-}^{10} DHHC_6}{dt} &\approx 0 \\
 \frac{dC_{2-}^{10} DHHC_6}{dt} &\approx 0 \\
 \frac{dC_{1-}^{11} APT}{dt} &\approx 0 \\
 \frac{dC_{2-}^{11} APT}{dt} &\approx 0
 \end{aligned}
 \tag{1.3.20}$$

and we derive the following system (1.3.21)

$$\begin{aligned}
 \frac{dC_{1-}^{00}DHH C_6}{dt} &= DHH C_6^{Tot} - \frac{K_f^{M.1.2} \cdot C_{1-}^{10}DHH C_6}{(C_{1-}^{10} - C_{1-}^{10}DHH C_6 - C_{1-}^{10}APT)} - C_{2-}^{00}DHH C_6 - C_{1-}^{10}DHH C_6 - C_{1-}^{01}DHH C_6 \\
 \frac{dC_{2-}^{00}DHH C_6}{dt} &= DHH C_6^{Tot} - \frac{K_f^{M.2.2} \cdot C_{1-}^{01}DHH C_6}{(C_{1-}^{01} - C_{1-}^{01}DHH C_6 - C_{1-}^{01}APT)} - C_{2-}^{00}DHH C_6 - C_{1-}^{10}DHH C_6 - C_{1-}^{01}DHH C_6 \\
 \frac{dC_{1-}^{10}APT}{dt} &= APT^{Tot} - \frac{K_b^{M.1.2} \cdot C_{2-}^{11}APT}{(C_{1-}^{11} - C_{1-}^{11}APT - C_{2-}^{11}APT)} - C_{1-}^{01}APT - C_{1-}^{11}APT - C_{2-}^{11}APT \\
 \frac{dC_{2-}^{10}APT}{dt} &= APT^{Tot} - \frac{K_b^{M.2.2} \cdot C_{1-}^{11}APT}{(C_{1-}^{11} - C_{1-}^{11}APT - C_{2-}^{11}APT)} - C_{1-}^{10}APT - C_{1-}^{11}APT - C_{2-}^{11}APT \\
 \frac{dC_{1-}^{01}DHH C_6}{dt} &= DHH C_6^{Tot} - \frac{K_f^{M.1.1} \cdot C_{1-}^{10}DHH C_6}{(C_{1-}^{00} - C_{1-}^{10}DHH C_6 - C_{1-}^{01}DHH C_6)} - C_{1-}^{01}DHH C_6 - C_{2-}^{00}DHH C_6 - C_{1-}^{01} \\
 \frac{dC_{1-}^{01}DHH C_6}{dt} &= DHH C_6^{Tot} - \frac{K_f^{M.2.1} \cdot C_{1-}^{01}DHH C_6}{(C_{1-}^{00} - C_{1-}^{10}DHH C_6 - C_{1-}^{01}DHH C_6)} - C_{1-}^{01}DHH C_6 - C_{2-}^{00}DHH C_6 - C_{1-}^{01} \\
 \frac{dC_{1-}^{11}APT}{dt} &= APT^{Tot} - \frac{K_b^{M.1.1} \cdot C_{1-}^{10}APT}{(C_{1-}^{10} - C_{1-}^{10}DHH C_6 - C_{1-}^{10}APT)} - C_{1-}^{01}APT - C_{1-}^{10}APT - C_{2-}^{11}APT \\
 \frac{dC_{2-}^{11}APT}{dt} &= APT^{Tot} - \frac{K_b^{M.2.1} \cdot C_{1-}^{01}APT}{(C_{1-}^{01} - C_{1-}^{01}DHH C_6 - C_{1-}^{01}APT)} - C_{1-}^{10}APT - C_{1-}^{01}APT - C_{1-}^{11}APT
 \end{aligned} \tag{1.3.21}$$

It is possible to see that (1.3.21) is similar to (22a-b) in [75] and equations (1.3.13) and (1.3.14) shown above, but it includes additional terms. In order to derive the tQSSA for this system, special care must be taken as stated in [76]. The exact tQSSA for this model does not exist, because the “total substrate” defined in (1.3.18) involves some complexes that are not relevant when calculating one of the reaction rates [76]. Therefore, in order to derive tQSSA for this system we need to define the pool for each of these complexes and subtract it from the total substrates, as done in equations (21-24) of [76]. This approximation allows us to match the polynomial derived from (1.3.21) with the one in (1.3.15), which can then be solved as in [75] (equations 24, 25, 26). The resolution of the polynomial allows to derive the generalized formula for tQSSA ((27) in [75]), which we used to model our system. The same approach shown above has been successfully used to model different network of competitive reactions and double modification mechanism, included the double phosphorylation/dephosphorylation of MAPK [76], which is very similar to the model presented in this paper.

3.3 Development of an *In-silico* labelling technique for model calibration and analysis

One of the most important outcomes of this project is the development of an *in-silico* labelling technique. One of the reasons that motivated the design of this method was the need of being able to replicate the labelling experiments performed *in-vivo*. In fact *In-silico* labelling is at the base of parameter estimation, allowing to evaluate the goodness of a given set of parameters used for model simulation; the goodness of fit of a set is computed calculating the distance between an experimental curve obtained from *in-vivo* labelling experiments, and the output of the model simulating the same experiment *in-silico*. This technique has also been successfully used to suggest simple modifications of the experimental labelling protocol that helped to better characterize properties of the proteins under analysis.

In this section we show how it is possible to perform different types of labelling experiments *in-silico*. In order to better illustrate the method we show its application on the calnexin model, but the exact same procedure is applied to any model discussed in this thesis. First we review the types of *in-vivo* labelling used in this project:

1. **³⁵S labelling**: cells are treated with ³⁵S cysteine/methionine for 20 minutes. Then the labelling is removed and the abundance of labelled calnexin is measured at different time points. These experiments are useful to capture the dynamics of protein synthesis and degradation.
2. **³H labelling**: cells are treated with ³H palmitate for 2 hours. Then the labelling is removed and the abundance of radiolabelled palmitoylated calnexin is measured at different time points. The difference with respect to ³⁵S experiments lies in the fact that here the label is attached to the palmitate molecule and not on the protein. Therefore only the population of protein that

gets palmitoylated will be visible with this technique. This type of labelling is useful to observe the dynamics of palmitate incorporation/loss.

3. **SNAP labelling:** This labelling consists of the fusion of 182 residual polypeptides (called SNAP) with the protein of interest. The protein is then expressed in the cell until it reaches physiological levels, and at that point the SNAP can be specifically and covalently labelled with a fluorescent dye for a specific time. The fluorescent dye is then removed from the culture to observe the decay of the labelled protein. The advantage of this technique is that it is possible to activate the label at the desired time. Is therefore possible to wait that the tagged protein reaches physiological levels before activating the label. Unfortunately due to the size of the tag, in some cases the tagged peptide may not folds properly, leading to a bias in the dynamics of the protein under analysis.

These labelling techniques were used to label WT and mutant version of the same protein, in which one or more palmitoylation sites were knocked out. Analyses of the mutants help to better characterize the dynamic of palmitoylation/depalmitoylation of each sites, and to investigate the relationship between stability and palmitoylation of individual sites.

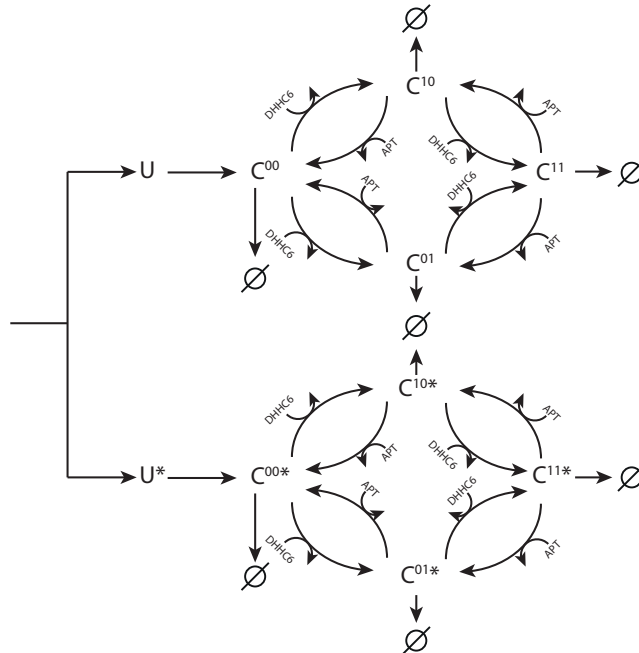
The *in-silico* labelling technique is based on the following assumption on *in-vivo* experiments:

1. When the labelling start the endogenous protein is in a steady state condition.
2. In ^{35}S labelling experiments, we add an excess of radiolabeled cysteines, which are much more abundant than the cysteines already present inside the cell. This implies that during the labelling phase only radiolabeled cysteines will be used to synthesize new proteins.
3. In ^3H labelling experiments we add an amount of radiolabeled palmitate which is much more abundant than the pool present inside the cell. This mean that during the labelling time only radiolabeled palmitate will be used to modify proteins.

- The pool of non-labelled protein is in competition with the tagged pool to get palmitoylated by a DHHC enzyme, and the same is true for the different palmitoylation sites present on a protein.

3.3.1 ^{35}S labelling

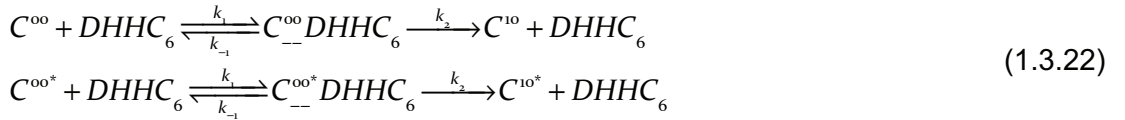
With this technique we introduce in the cell an excess of radiolabelled amino acids. Once the cell absorbs the label, the newly synthesized proteins will be tagged with radioactive amino acid ending up with pools of tagged proteins. Labelled proteins can then be purified and quantified. One of the first steps to replicate this experiment *in-silico* is to be able to distinguish in the model between the labelled and non-labelled pools of protein. This is done by taking the calnexin model in Figure 3.2, and duplicating its species as in **Error! Reference source not found.**



Error! Reference source not found. shows two identical but separated pools of protein. One pool represents the WT endogenous calnexin, while the other serve to

simulate the labelled protein that will be synthesized in the cell once the labelling is active (marked with * in **Error! Reference source not found.**). There is only one synthesis term, in agreement with assumption n.2 described above. Synthesis can be switched between endogenous or labelling species depending from which part of the experiment we want to simulate. During a labelling experiment *in-vivo*, after introducing the label in the cell we will have two distinct populations. One is the endogenous protein, and the other represents the newly synthesized proteins that have been labelled. The presence of two distinct populations will affect their kinetics of palmitoylation. This is because the two populations compete in order to get palmitoylated by the same enzyme, since the only difference between the two pools of protein is the presence/absence of the label. During *In-silico* experiments this competition behaviour is captured by additional competitive terms in each reaction rate. These terms quantify the competition between labelled and non-labelled proteins for each enzymatic reaction. Competition between the different species of the model have been implemented the same way as in [75]. Below we show the steps needed to implement competition for *in-silico* labelling:

If we consider the competitive reaction scheme (* indicates labelled species):



We can formulate the following mass action equations:

$$\begin{aligned}
 \frac{dC^{oo}}{dt} &= k_{-1} \cdot C^{oo}DHHC_6 - k_1 \cdot C^{oo} \cdot DHHC_6, \\
 \frac{dC^{oo}DHHC_6}{dt} &= k_1 \cdot C^{oo} \cdot DHHC_6 - (k_{-1} + k_2) \cdot C^{oo}DHHC_6, \\
 \frac{dC^{oo*}}{dt} &= k_{-1} \cdot C^{oo*}DHHC_6 - k_1 \cdot C^{oo*} \cdot DHHC_6, \\
 \frac{dC^{oo*}DHHC_6}{dt} &= k_1 \cdot C^{oo*} \cdot DHHC_6 - (k_{-1} + k_2) \cdot C^{oo*}DHHC_6.
 \end{aligned}
 \tag{1.3.23}$$

Along with the conservation law:

$$DHHC_6^{Tot} = DHHC_6 + C_{--}^{oo*} DHHC_6 + C_{--}^{oo} DHHC_6 \quad (1.3.24)$$

If we set:

$$\begin{aligned} \frac{dC_{--}^{oo} DHHC_6}{dt} &= 0, \\ \frac{dC_{--}^{oo*} DHHC_6}{dt} &= 0. \end{aligned} \quad (1.3.25)$$

And we introduce the new quantities:

$$\begin{aligned} \overline{C^{oo}} &= C^{oo} + C_{--}^{oo} DHHC_6, \\ \overline{C^{oo*}} &= C^{oo*} + C_{--}^{oo*} DHHC_6. \end{aligned} \quad (1.3.26)$$

We can re-write equations in (1.3.23) using tQSSA and implementing competition among the different substrates:

$$\begin{aligned} \frac{d\overline{C^{oo}}}{dt} &= -V_{\max} \cdot \frac{\overline{C^{oo}}}{K_M \cdot \left(1 + \frac{\overline{C^{oo*}}}{K_M^*}\right) + \overline{C^{oo*}} + DHHC_6}, \\ \frac{d\overline{C^{oo*}}}{dt} &= -V_{\max} \cdot \frac{\overline{C^{oo*}}}{K_M^* \cdot \left(1 + \frac{\overline{C^{oo}}}{K_M}\right) + \overline{C^{oo}} + DHHC_6}. \end{aligned} \quad (1.3.27)$$

Where $V_{\max} = k_2 \cdot DHHC_6^{Tot}$ and $K_M = \frac{k_{-1} + k_2}{k_1}$.

Once we re-write our model according to the previous kinetic terms we proceed as follows (Figure 3.4):

1. We simulate the model with the synthesis opened towards endogenous calnexin until it reaches steady state.

2. When the system is in steady state we switch the synthesis towards the labelled protein for an amount of time that matches with the experimental labelling time.
3. At the end of the labelling time we redirect the synthesis from the labelled system back to the endogenous protein and we monitor the concentration of tagged protein at the same time points as the experiment, with time zero (t=0) being the end of labelling time.

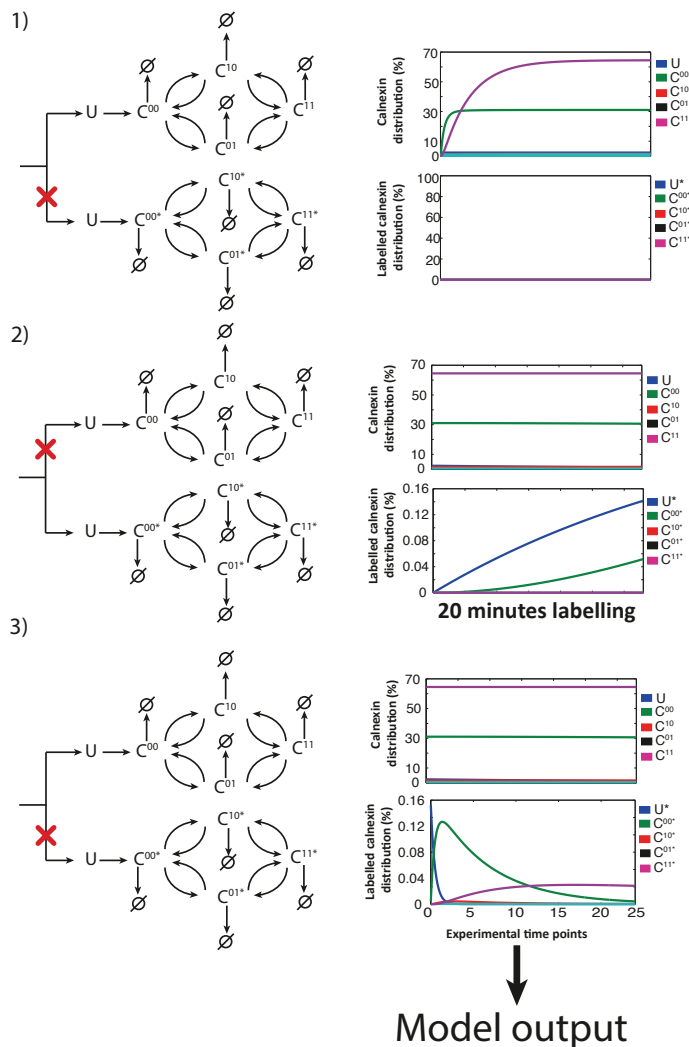


Figure 3.4. Different phases of a ^{35}S *in-silico* labelling experiment

3.3.2 ^3H labeling

This type of labelling consists in the addition of an excess of radiolabelled palmitate to cells. During the labelling phase the cell will incorporate radiolabelled palmitate, therefore all the proteins undergoing palmitoylation will be tagged with labelled palmitate. The pool of tagged protein of interest can then be tracked and quantified as in ^{35}S labelling. The difference from ^{35}S labelling is that the protein is tagged after acquiring the palmitate and not during the synthesis. As consequence only palmitoylated protein will be labelled (Figure 3.5). If only a fraction of a protein pool is palmitoylated, the part that is not palmitoylated won't be detected by ^3H labelling.

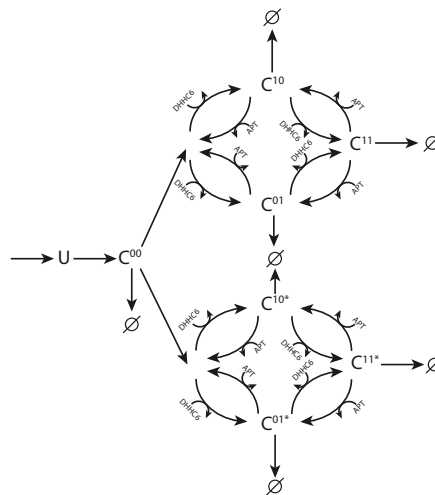
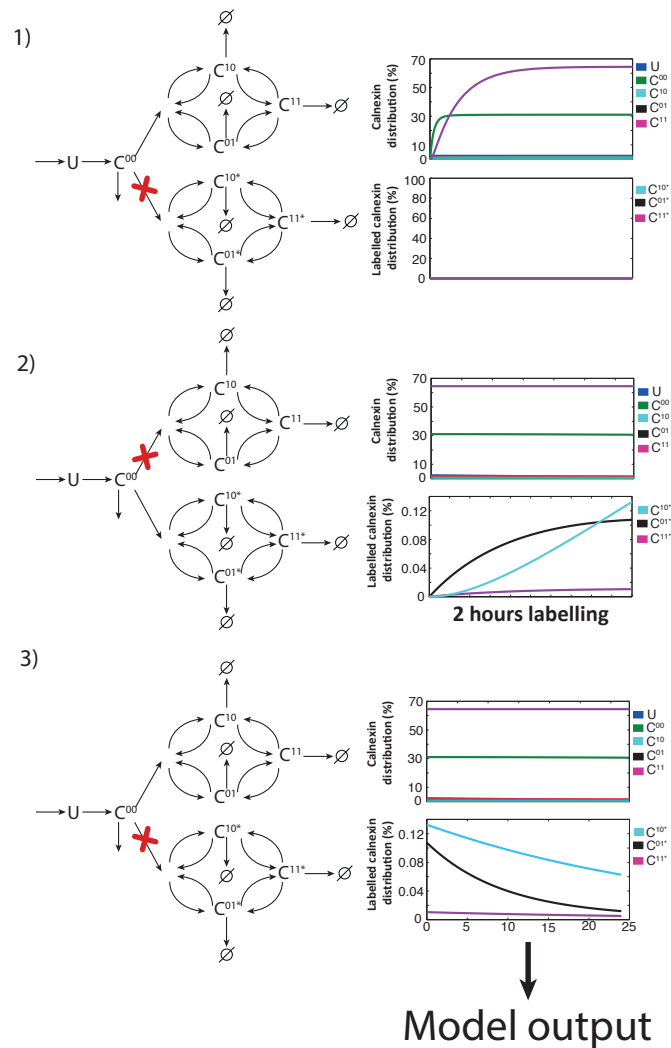


Figure 3.5. Network topology of the calnexin model used for ^3H *in-silico* labelling.

The modifications that need to be done to the model to include this type of labelling are similar with the ones used for ^{35}S labelling. As before the model is duplicated in size, one pool representing endogenous species and the other the labelled species, interacting with each other thanks to the competitive terms in the reaction rates (see 3.3.2). With palmitate being the labelling molecule, the pool of tagged calnexin will contain two less species. These are the species that are not palmitoylated (i.e. U and C^{00} , as in Figure 3.5). Simulations are performed following the same procedure as for ^{35}S (Figure 3.6), taking care of adjusting the labelling time to match the experimental time for ^3H -palmitate labelling (2 hours).

Figure 3.6. Different phases of a ^3H *in-silico* labelling experiment

3.3.3 Model parameterization with heuristic optimization

At the time of writing very little kinetic data are available on the enzymatic process of palmitoylation. No kinetic information is available about the 23 enzymes that catalyse palmitoylation in human. Due to the lack of data, all the parameters of the different models were estimated using a genetic algorithm (GA). GA is a heuristic global optimization algorithm that mimics the process of natural selection [79]. In order to

estimate the parameters, multiple datasets coming from experimental results were considered. Time course labelling experiments were performed in order to characterize the dynamics of calnexin synthesis/degradation and incorporation/loss of palmitate (see chapter 3.3). The procedure we used for parameter optimization follows a classic approach where the experimental data for a certain protein are divided in two groups. The first dataset is used to calculate the goodness of fit of each set of parameter generated by the GA, this subset of data is called calibration set. The second subset is used to validate the output of the model and test its prediction capabilities on data that were not used for parameter estimation. The genetic algorithm we used is implemented in MATLAB and it is named *gamultiobj*. This algorithm finds the overall minimum of multiple objective functions by minimizing the difference between the output of the model and the experimental data of the calibration set. A complete description of the genetic algorithm can be found in [78]. Here we report the main steps:

1. The algorithm begins by creating a random initial population set of parameters. Each set of the population is a set of all the modelling parameters needed to perform the simulations. In order to generate the initial population we have to specify an upper and lower bound for the parameters values, the GA would then generate an initial population by uniformly sampling random values in between the bounds. The values of the bounds are mainly dependent from the concentration of the protein under analysis. For example, since calnexin has a concentration in the order of magnitude of μM we set bounds for the parameter that span over the estimated calnexin concentration (from $10^{-3} \mu M$ to $10^3 \mu M$).
2. At each step, the algorithm uses the sets in the current iteration to create the next population. To create the new population sets, the algorithm performs the following steps:
 - a. Scores each set of the current population by computing its fitness value.
 - b. Selects sets, based on their fitness.

- c. Produces sets for the next population from the sets of the current population. The new sets are produced either by making random changes to the parameters of a single set (mutation) or by combining the vector entries of a pair of parameter sets (crossover).
3. Replaces the current population with the new sets obtained to form the next generation.
4. The algorithm stops when one of the stopping criteria is met. These criteria are:
 - a. The average relative change in the best fitness function value is less than the fitness function tolerance (10^{-6}).
 - b. The number of iterations reaches the value of 10^4 .

Once GA has converged to a set of solutions we repeat the optimization process iteratively, adjusting each time the bounds on the parameter. This process is executed until reaching a satisfactory match between simulated data and experiments. This procedure was applied to all the parameters except for the degradation rate constant of the different species. In fact this parameter can be estimated from the turnover rate $t_{1/2}$ of the different mutants measured experimentally as:

$$k_d = \frac{\log(2)}{t_{1/2}} \quad (1.3.28)$$

The values found for the degradation rate constants were allowed to vary $\pm 20\%$ during parameter estimation. We have also considered the experimental errors based on the standard deviation values of the experimental measurements. We should also mention that none of the final estimated parameter values was at the upper or lower bounds used for them during the GA estimation procedure

Genetic algorithms, as many others optimization methods, are based on the minimization of a cost function. The cost function is computed for each set of parameters generated by the GA, to evaluate its goodness. Each set is used to simulate the experiments of the calibration set, and the output of the model is then compared against the experimental data as follow:

$$f(x_i) = \begin{bmatrix} f_1(x_i) \\ f_2(x_i) \\ \vdots \\ f_k(x_i) \end{bmatrix}, \quad f_k(x_i) = \sum_{j=1}^{N_m} (x_j^{\text{exp}} - x_j^{\text{sim}})^2, \quad k = [1 \dots N_{\text{obj}}] \quad i = [1 \dots N_{\text{ind}}] \quad (1.3.29)$$

Where $f(x_i)$ is the vector of fitness function values for the set of parameters x_i , $f_k(x_i)$ is the value of fitness function evaluated for a single experimental result in the calibration set. For each set of parameter x_i the GA computes the fitness value $f_k(x_i)$ of each experiment in the calibration set as the sum of the squared difference between the experimental measurement x_j^{exp} and the output of the model x_j^{sim} for each experimental point in a single experiment. N_m is the number of measurements that were taken in a single labelling experiment, k is the number of objectives we want to minimize and N_{ind} is the number of individuals in the GA population. Because of the presence of multiple objectives there does not exist a single solution that simultaneously optimizes each objective, so the algorithm provides as output a local Pareto set of solutions, which are equally optimal with respect to the fitness function we defined. The solutions in the Pareto set are equally optimal in the sense that for each of them none of the objective functions $f_k(x_i)$, can be improved in value without deteriorating the quality of the fitness in some of the other objective values. In order to be more accurate and to reduce the variability in the output of the model, from the Pareto set provided by the GA, we selected those set of parameter that fitted best the calibration data. Selection of this subset was done as follow: for each optimized set of parameters, the GA provides the corresponding score for each objective (we have one objective for each experiment used as calibration). The first step is to scale the scores of each objective with respect to the maximum scores obtained for that particular objective. This is done for each parameter set in the Pareto (1.3.30):

$$f(x_i) = \begin{bmatrix} f_1(x_i) \\ f_2(x_i) \\ \vdots \\ f_k(x_i) \end{bmatrix} \rightarrow \overline{f(x_i)} = \begin{bmatrix} \frac{f_1(x_i)}{\max(f_1)} \\ \frac{f_2(x_i)}{\max(f_2)} \\ \vdots \\ \frac{f_k(x_i)}{\max(f_k)} \end{bmatrix} \quad (1.3.30)$$

For each parameter set in the Pareto we then compute the sum of each fitness value to obtain a single value representing the overall goodness of fit of a set of parameters with respect to all the objectives (1.3.31). This value was then scaled to be in the range between 0 and 1 (1.3.32). The sets of parameters with a fitness value of 0.3 or less were selected in the final set of parameters used for simulations and predictions.

$$f(x_i)_{Tot} = \sum_{j=1}^{N_{obj}} f_j(x_i) \quad (1.3.31)$$

$$\overline{f(x_i)_{Tot}} = \frac{f(x_i)_{Tot}}{\max(f_{Tot})} \quad (1.3.32)$$

3.3.4 Studying impact of modelling parameters with sensitivity analysis

3.3.4.1 Local Sensitivity analysis:

After each model calibration a common practice is to perform parametric sensitivity analysis (SA), with the aim of gaining more insight on how the various parameters of the model affect its output. SA was performed in the initial part of the project using a

simple One-at-a-time (OAT) method, which consisted of changing one parameter at a time and observing its effect on the output of the model.

We explain here the step-by-step procedure adopted to calculate sensitivity for each of the model considered in this project. The description refers to the calculation of sensitivity indexes in the case of calnexin's half-life, but the same procedure applies when calculating sensitivity indexes with respect to others properties or to other models, like for example total protein abundance or total level of palmitoylated protein:

1. For each set of parameters from the population of GA optimized sets we simulate a ^{35}S labelling experiment (See 3.3). In this type of experiments we label calnexin with radioactive amino acids for 20 minutes, we then stop the labelling and measure the abundance of radiolabelled calnexin in the cell at different time points. The half-life of calnexin correspond to the time point at which its concentration drop to 50% of the concentration measured at the end of the labelling phase. The values measured this way will be taken as reference points.
2. Iteratively, each parameter of the set is increased by 1% of its value. This is done for each optimized set.
3. After perturbation of a parameter we repeat the process from point 1. The values measured this way will be stored as perturbed points.
4. The value of the parameter modified is restored to its original value

Points 2 to 4 are repeated for each parameter of a set, resulting in N_p different perturbed simulations, were N_p is the number of parameter needed to simulate the model.

5. In order to calculate dynamic sensitivity analysis, a sensitivity index for each parameter and each time step of a simulation are calculated as follows:

$$S_i(t_j) = \frac{f(P_{x_i^*}, t_j) - f(P_r, t_j)}{x_i^* - x_i^r} \cdot \frac{x_i^r}{f(P_r, t_j)}, \quad i = [1 \dots N_p] \quad j = [t_1 \dots t_{end}] \quad (1.3.33)$$

Where $S_i(t_j)$ is the sensitivity index of the i^{th} parameter at the time step j . $f(P_{x_i^*}, t_j)$ is the model output at time t_j for the set of parameter P in which the parameter x_i was varied of 1%. $f(P_r, t_j)$ is the model output at time t_j for the set of parameter P_r used as reference. x_i^* is the i^{th} perturbed parameter while x_i^r is the same parameter but with reference value. N_p is the number of parameters of a set.

3.3.4.2 Global sensitivity analysis

This section has been designed and wrote with the help of Robin Denhardt Eriksson

Although local SA has been used extensively in this project, its limitations led to a change in the approach to SA calculation in the final part of the project. The main problems with local SA are that sensitivity indexes are always calculated with respect to a reference point (i.e. a particular steady state). If the reference point change, sensitivity indexes need to be recalculated on the base of the new reference point. Therefore local SA does not consider parameters over an entire range, but sensitivity indexes are calculated in a small region around a reference point. Another important limitations of local SA is that during the calculation of local sensitivity indexes, since the Finite Difference Method is used to approximate the derivative of the function under analysis, we need to vary our input parameters of a sufficiently small quantity in order for the approximation to hold. This implies that local sensitivity indexes are only valid for small deviation from the reference point. Moreover depending on the values of the parameters considered, the variation in the parameter need to be adjusted to be sufficiently small.

To avoid these problems and improve the quality of our analysis we decided to use a global method to estimate sensitivity indexes. These methods are also called variance-based since they analyse how the variance in the input parameters propagate to variance in the model output [81]. Global sensitivity overcome the problems of the local method cited before and was therefore chosen to substitute local sensitivity in the final part of the project. In this chapter we are going to briefly overview how variance-based methods work.

Let f be a square-integrable function depending on k parameters x_i , and with output Y :

$$Y = f(x_1, x_2, \dots, x_k) \quad (1.3.34)$$

Since f is square-integrable we can expand the function in terms of increasing dimensionality:

$$f(\mathbf{x}) = f_0 + \sum_{i=1}^n f_i(x_i) + \sum_{1 \leq i < j \leq n}^k f_{ij}(x_i, x_j) + f_{12\dots n}(x_1, x_2, \dots, x_n) \quad (1.3.35)$$

$$\mathbf{x} = (x_1, x_2, \dots, x_n)$$

This kind of expansion is called High Dimensional Model Representation (HDMR), it is finite and contains $2^k - 1$ terms [81,82]. In (1.3.35) the term f_0 correspond to the mean of the output $f(\mathbf{x})$. The sum of terms f_i corresponds to additive effects of single input parameters on the model output, while the terms of higher order correspond to interactions between parameters. It is proved that if all the terms of the decomposition, apart from f_0 , have a zero mean then these terms are orthogonal in pairs and conditional expected values may be used to compute them [81]. Therefore, to calculate the analysis of variance (ANOVA) from (1.3.35), we just need to consider the variance on each side of the equations [81]:

$$\text{Var}(f(\mathbf{x})) = \sum_{i=1}^n \text{Var}(f_i(x_i)) + \sum_{1 \leq i < j \leq n}^k \text{Var}(f_{ij}(x_i, x_j)) + \text{Var}(f_{12\dots n}(x_1, x_2, \dots, x_n)) \quad (1.3.36)$$

Sobol indexes are obtained by normalization of each term in (1.3.36) by the total variance $\text{Var}(f(\mathbf{x}))$ of the model output:

$$S_i = \frac{\text{Var}(f_i)}{\text{Var}(Y)} = \frac{\text{Var}_{X_i}(E_{X_{-i}}(Y | X_i))}{\text{Var}(Y)} \quad (1.3.37)$$

In (1.3.37) the notation $E_{X_{-i}}$ means that the model output Y is averaged across all values of X_{-i} , all parameters other than X_i . The notation Var_{X_i} means that the variance is calculated across all values of X_i . The indexes calculated in (1.3.37) are known as “first order” sensitivity indexes. The value of a first order sensitivity index represents the fraction of variance in the output that is attributed by the variance in the input parameter x_i . This index captures the sensitivity of the model towards x_i without taking into account interactions effects with the other parameters. The interaction effects can be captured measuring the variance of higher order terms in (1.3.36). Since a model can have many parameters, and therefore many higher order terms, the computation time required to calculate all the higher order term may be prohibitive. To avoid the calculation of all the possible interactions effects, which increase exponentially with the numbers of parameters, Sobol introduced what is called total effect [81]. For each parameter the Total effect is the sum of the first effect and all the higher order effects:

$$S_{t,i} = S_i + \frac{\sum_{j>i}^k \text{Var}(f_{ij}(x_i, x_j))}{\text{Var}(f(\mathbf{x}))} + \frac{\sum_{j>i}^k \sum_{l>i}^k \text{Var}(f_{ijl}(x_i, x_j, x_l))}{\text{Var}(f(\mathbf{x}))} + \dots + \frac{\text{Var}(f_{12\dots k}(x_1, x_2, \dots, x_k))}{\text{Var}(f(\mathbf{x}))} \quad (1.3.38)$$

The total effect of X_i can be considered as the total variance of the model output minus all other 1st order effects and interactions between parameters excluding X_i , all of this normalised with the output variance, greatly simplifying the calculation [81]:

$$S_t = \frac{V(f(\mathbf{x})) - V_{x_i}(E_{x_i}(f(\mathbf{x})|x_{-i}))}{V(f(\mathbf{x}))} \quad (1.3.39)$$

This procedure greatly reduces the computation time, especially with models that have many parameters, since we don't have to compute each interaction term individually. The value $S_t - S_i$ allows to estimate the amount of variability in the output of the model that can be attributed to interaction effects of the parameter x_i with the other parameters [81].

Here we illustrate how we implemented the algorithm for calculation of global sensitivity. The methods is accurately described in [81,83], here we report the main steps.

The algorithm start by generating three matrices of size (N, k) , A, B and C_i (1.3.40), where N is the number of samples we generate, and k is the number of parameter needed to successfully simulate the model. For each matrix every row corresponds to a parameter set, which contain all the parameters needed to simulate a model. Each column contains different sampled values of a particular parameter x_i .

$$A = \begin{pmatrix} a_1^1 & \cdots & a_k^1 \\ \vdots & \ddots & \vdots \\ a_1^N & \cdots & a_k^N \end{pmatrix} \quad B = \begin{pmatrix} b_1^1 & \cdots & b_k^1 \\ \vdots & \ddots & \vdots \\ b_1^N & \cdots & b_k^N \end{pmatrix} \quad (1.3.40)$$

$$C = \begin{pmatrix} b_1^1 & \cdots & a_1^1 & \cdots & b_k^1 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ b_1^N & \cdots & a_1^N & \cdots & b_k^N \end{pmatrix}$$

While A and B are independently generated matrices, C_i is composed by all the columns of B , except column i that is taken from A . In the end we will generate k C matrices. The model outputs are then evaluated and stored in vectors of length N :

$$y_A = f(A), \quad y_B = f(B), \quad y_{C_i} = f(A = C_i) \quad (1.3.41)$$

The 1st order and total effect sensitivity indices are then calculated in the following way (1.3.42):

$$S_i = \frac{\frac{1}{N} \sum_{j=1}^N y_A^j \cdot y_{C_i}^j - f_o^2}{\frac{1}{N} \sum_{j=1}^N (y_A^j)^2 - f_o^2} \quad (1.3.42)$$

$$S_t = 1 - \frac{\frac{1}{N} \sum_{j=1}^N y_B^j \cdot y_{C_i}^j - f_o^2}{\frac{1}{N} \sum_{j=1}^N (y_A^j)^2 - f_o^2}$$

Where f_o is the mean of the model output for each set of parameters (1.3.43):

$$f_o = \frac{1}{N} \cdot \sum_{j=1}^N y_A^j \quad (1.3.43)$$

3.4 Stochastic simulation as a tool for molecule tracking

During this project we were interested in determining different properties characteristic of single protein molecules. For example interesting features can be the average palmitoylation time of a single molecule of calnexin or the time that DHHC₆ spend in the different palmitoylation states. This kind of properties are not easily accessible when working with deterministic models, mainly due to the fact that we can't observe single molecules, but we have instead molecules fluxes that flow from one state to another. To overcome this problem we developed a technique that allows to perform *in-silico* single molecule tracking. This method have been developed on the basis of the *in-silico* labelling technique described in chapter 3.3. In order to be able to track a single molecule inside a model we designed a stochastic versions of the deterministic models developed in this thesis project. The structure of the models remains the same as described in chapter 3.3 for ³⁵S labelling (**Error! Reference**

source not found.) We always have the presence of two sets of identical species. What differ from the *in-silico* labelling technique are the initial conditions for the pool of labelled protein. For single molecule tracking in the labelled pool we put just one molecule of the protein under analysis in the unfolded state; this will be the tracked protein. The synthesis term is always directed towards synthesis of endogenous protein, to ensure that in the pool of labelled protein we will always have only one protein. As for the labelling technique the two set of species (labelled and not labelled) are linked through the competition terms for enzymes substrates (see chapter 3.3.1). The single labelled protein will feel the presence of the entire pool of endogenous molecules; therefore its dynamics will behave like the protein is in steady state conditions. After setting up, we simulate the model until the protein in the labelled pool is degraded, and for each time step of the simulation we record time and position of the protein in the system. An illustration of the procedure is visible in Figure 3.7. The analysis of the single labelled molecule route allows to determine different properties of the protein. Due to the stochastic nature of the simulations the procedure described above is repeated from 5000 to 10000 times depending on the protein under analysis. This ensures that stochastic simulations have converged and therefore the calculated properties are reliable.

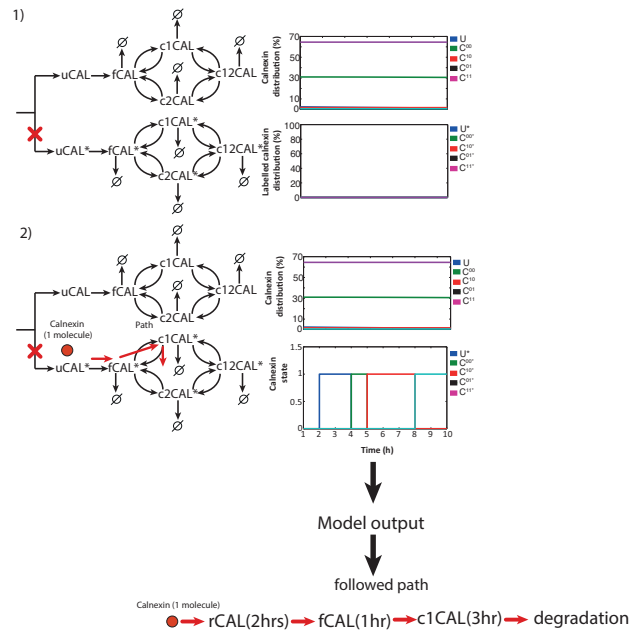


Figure 3.7. Different phases of a single molecule tracking experiment.

3.5 Stochastic model formulation:

To perform single molecules tracking we created stochastic versions of our deterministic models. One of the main challenges during the development was that some of the proteins studied in this project are very abundant in the cell, like for example calnexin. This protein is present in HeLa cell in a number of copies between $5 \cdot 10^5$ and 10^6 [21,40,41]. Since we want to measure a property of the system in steady state conditions, we have to perform simulations with a total number of molecules that match with steady states values. The high number of molecules present in steady state for some of our models made stochastic simulations computationally expensive. Therefore, in order to perform simulations in an acceptable amount of time and resources we used the same reduction approach as we did in the deterministic models. We performed stochastic simulations using a model described with tQSSA. Different studies have shown that tQSSA is an approximation that can be applied to stochastic models without losing accuracy [84,85]. Validation results of this method have already been obtained for the Goldbeter switch, MAPK cascade, Michaelis-Menten kinetics

and others kinetic models [86]. Therefore, the models that we use are the same as in the deterministic simulations, but as required by the stochastic simulations, model parameters were converted from concentrations to number of molecules (see below). The stochastic simulations were performed using the first reaction method, a variant of the Gillespie algorithm [87].

3.6 Conversion of deterministic parameters to stochastic:

Stochastic simulations can't be done with parameters that were estimated for deterministic simulations. Deterministic parameters need to be converted; this transformation involves a change of units, from concentration to number of molecules. The following assumption are needed in order to successfully convert the parameters:

- The volume of HeLa cells is estimated to be $1540 \mu M$ on average (from bionumbers database).
- We estimated the ER volume to be approximately 3% of the average cell volume. Therefore the ER volume is equal to: $V_{ER} = 48.5 \mu M = 4.85 \cdot 10^{-14} L$

The first step for conversion is to reduce the sets of parameters estimated from the GA to a single set. While we could derive the equivalent stochastic models for each set of parameter for a given protein, this would further require at least 5000 simulations for each one of the sets of parameter and it would have been computationally too expensive (we currently perform stochastic simulation of one set of parameters in 2 days).

In order to derive a single set we averaged each sets of parameter for a given protein. To ensure that this set of parameter was still able to maintain the prediction capabilities shown by the original group of parameter sets we plotted its output against the experimental dataset used for calibration and validation (Figure 3.8).

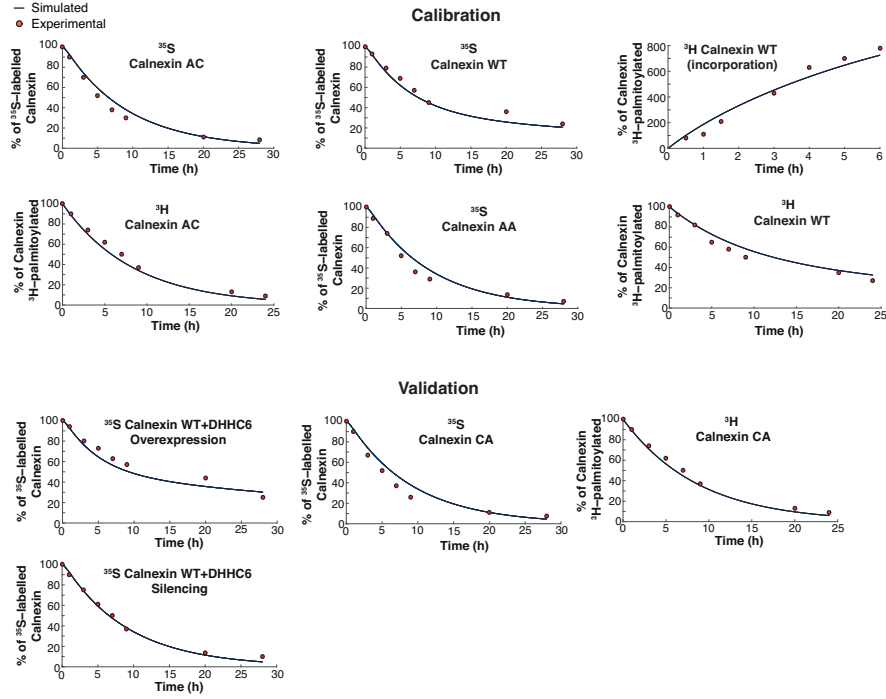


Figure 3.8. Results of model calibration and validation for a single set of parameters obtained by averaging GA output.

The conversion of the parameter proceed then as follow:

- Parameter with units $\frac{\mu M}{h}$:

$$P_i \left(\frac{\mu M}{h} \right) = P_i \cdot 10^{-6} \cdot \left(\frac{M}{h} \right) = P_i \cdot 10^{-6} \cdot \left(\frac{M}{h} \right) \cdot V_{ER} (L) \cdot N_A \left(\frac{\text{molecules}}{\text{mole}} \right) = P_i \cdot 2.92 \cdot 10^4 \frac{\text{molecules}}{h}$$

- Parameter with units μM

$$P_i (\mu M) = P_i \cdot 10^{-6} \cdot (M) = P_i \cdot 10^{-6} \cdot (M) \cdot V_{ER} (L) \cdot N_A \left(\frac{\text{molecules}}{\text{mole}} \right) = P_i \cdot 2.92 \cdot 10^4 \text{ molecules}$$

- No conversion is needed for parameter with units $\frac{1}{h}$

Chapter 4 - Model-Driven Understanding of Palmitoylation Dynamics: Regulated Acylation of The Endoplasmic Reticulum Chaperone Calnexin

This Chapter is published with the following details:

Dallavilla T, Abrami L, Sandoz PA, Savoglidis G, Hatzimanikatis V, van der Goot FG (2016) *Model-Driven Understanding of Palmitoylation Dynamics: Regulated Acylation of the Endoplasmic Reticulum Chaperone Calnexin*. *PLoS Comput Biol* 12(2): e1004774. doi:10.1371/journal.pcbi.1004774

4.1 Abstract

Cellular functions are largely regulated by reversible post-translational modifications of proteins, which act as switches. Amongst these, S-palmitoylation is unique in that it confers hydrophobicity. Due to technical difficulties, the understanding of this modification has lagged behind. To investigate principles underlying dynamics and regulation of palmitoylation, we have here studied a key cellular protein, the ER chaperone calnexin, which requires dual palmitoylation for function. Apprehending the complex inter-conversion between single-, double- and non- palmitoylated species required combining experimental determination of kinetic parameters with extensive mathematical modelling. We found that calnexin, due to the presence of two cooperative sites, becomes stably acylated, which not only confers function but also a remarkable increase in stability. Unexpectedly, stochastic simulations revealed that palmitoylation does not occur soon after synthesis, but many hours later. This prediction guided us to find that phosphorylation actively delays calnexin

palmitoylation in resting cells. Altogether this study reveals that cells synthesize 5 times more calnexin than needed under resting condition, most of which is degraded. This unused pool can be mobilized by preventing phosphorylation or increasing the activity of the palmitoyltransferase DHHC6.

4.2 Author summary

The endoplasmic reticulum (ER) is the largest intracellular organelle of mammalian cells. It is responsible for many fundamental cellular functions, such as folding, quality control of membrane and secreted protein, lipid biosynthesis, control of apoptosis and calcium storage. Recent studies have shown that many ER membrane proteins are lipid modified. We therefore hypothesized that palmitoyltransferases, the enzymes responsible for this modifications, act as a regulator of the mammalian ER, controlling the function of a network of key proteins through reversible acylation. In this work we combine computational methods with experimental determination of parameters to study the mechanisms and properties of ER palmitoylation, using as a model the palmitoylation of the ER protein calnexin. The systematic analysis of the mathematical model, built and calibrated with the help of experimental data, shows that Calnexin palmitoylation leads to a 9-fold increase in half-life and that a long delay separates synthesis from palmitoylation in unstimulated cells. Surprisingly during this delay, 75% of synthesized calnexin is degraded before being palmitoylated. We hypothesize that this unexpected apparent inefficiency is a design principle that provides the cell with a means to post-translationally tune the calnexin content.

4.3 Introduction

Reversible post-translational modifications of proteins allow cells to regulate processes in time and in space [25–29]. Amongst these, S-palmitoylation is unique in that it confers hydrophobicity to proteins by covalent attachment of a fatty acid chain to cysteine residues [15,32,47]. In the cytoplasm, this enzymatic reaction is mediated

by palmitoyltransferases of the DHHC family and reversed by acyl protein thioesterases (APTs) [15,32,44]. Recent large-scale palmitoyl-proteome profiling studies have jointly revealed that hundreds of proteins, with major cellular functions, undergo this lipid modification in mammalian cells [16,18,21,40,41]. Although S-palmitoylation was identified more than 30 years ago, our understanding of this modification, its dynamics, its regulation and its consequences on protein properties, is still rudimentary.

The aim of this paper is to study the palmitoylation events, and their dynamics, occurring on a key component of the endoplasmic reticulum (ER), the type I transmembrane protein calnexin. Here we report the step-by-step design and output analysis of the first model of a palmitoylation network. Besides studying palmitoylation, another valuable objective of this work has been the estimation of system parameters which cannot be estimated by simple experiments, such as the time required for calnexin to get double palmitoylated or the half life of the palmitoylated species, but instead, they require the consideration of the system as a whole.

Calnexin is best known for its function as a lectin-like chaperone involved in the folding of glycosylated proteins in the lumen of the ER [88]. It is also involved in regulating calcium homeostasis at ER-mitochondria contact sites [89]. More recently, we have found that calnexin can act as an ER sensor, modulating the transcriptional response of cells to EGF in an ER-stress dependent manner [90]. Importantly, the ability of calnexin to assist folding of newly synthesized proteins, to control calcium signalling and to modulate the EGF signalling response, all require its palmitoylation [34,87–89].

Calnexin is composed of a large well-folded N-terminal luminal domain that carries the chaperone activity. It is followed by a single transmembrane domain that terminates with two cysteine residues at positions 502 and 503, which are the sites of palmitoylation [34,91]. Even though the ER contains numerous DHHC enzymes [56], calnexin palmitoylation is mediated exclusively by DHHC6 [34]. The palmitoylation

sites are followed by a 90 residue cytosolic tail that is predicted to be disordered and contains multiple phosphorylation sites [92]. This cytosolic domain has multiple functions. It allows association of calnexin with the ribosome translocon complex [91,93] but can also be proteolytically released following specific stimuli such as apoptotic drugs [94] or EGF [90]. In the latter case, the released cytosolic tail binds to PIAS3 (Protein Inhibitor of Activated STAT) and thereby promotes EGF-induced STAT3-mediated transcriptional response to EGF [90].

Since calnexin has two sites of palmitoylation, it can exist in cells under different forms: non-modified, palmitoylated on one site or the other, or on both. To grasp the full complexity of the system, we combined mathematical modelling of the system with experimental determination of the kinetics of palmitate acquisition and turnover, and of protein degradation for wild type (WT) and palmitoylation-deficient mutants. We set up a general model consisting of ordinary differential equations (ODE), describing the dynamic transitions between states/species of the network. We defined the topology of our network as a combination of well-known inter-convertible cycles described by Goldbeter and Koshland [72]. Combinations of such subunits have been used to successfully describe biological systems in which proteins undergo multiple covalent modifications, especially in the context of protein kinase signalling networks [69,95,96]. Since palmitoylation has, as phosphorylation, the potential to control protein function in a switch-like manner, a similar approach appeared suitable to model calnexin palmitoylation.

Mathematical modelling allowed us to access an unprecedented level of understanding of the dynamics and the complexity of inter-convertible species of the same protein undergoing various post-translational modifications on multiple sites and access key parameters that are not directly measurable through experimentation. We could in particular estimate the half-life of single or dual palmitoylated calnexin, the off-rate of palmitate from a specific site upon occupancy of the other site or the time that separates calnexin synthesis from palmitoylation events, all together control the cellular levels and thus activity of this key ER chaperone.

4.4 Results

4.4.1 Mathematical model of calnexin palmitoylation

Calnexin was recently shown to rely on palmitoylation to perform its major functions [34,89–91]. This raises the question as to which percentage of the total calnexin population is at a given time palmitoylated. At present, no reliable method enables to determine the percentage of a protein that is palmitoylated and to differentiate single from double palmitoylation. To estimate the species distribution and understand the dynamics of the inter-conversion between them, we therefore developed a mathematical model of the calnexin palmitoylation cycle. Modelling was performed as an open system, including protein synthesis, and degradation of all species (Figure 4.1A, Table S4.2)

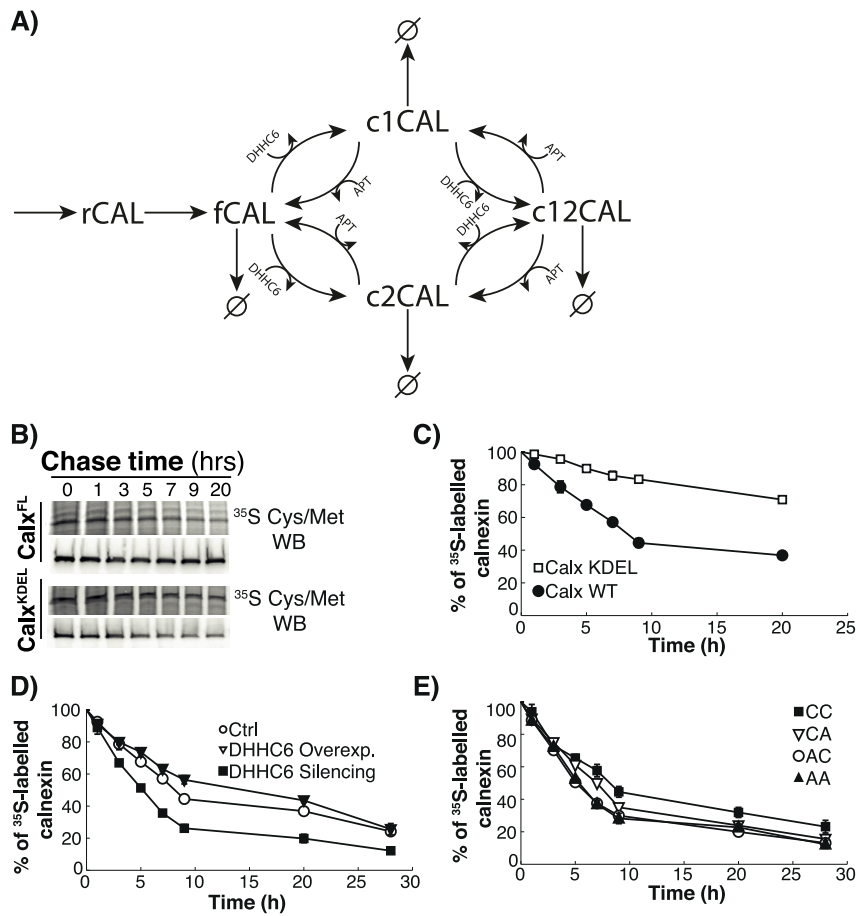


Figure 4.1. Calnexin palmitoylation model and kinetics of decay. (A) Calnexin is first synthesized (rCAL). During the folding process the protein assumes the proper conformation (fCAL), which then can be palmitoylated twice by DHHC6. The first palmitoylation can occur on either of the two sites leading to c1CAL or c2CAL. The second palmitoylation events leads to c12CAL. We assume that both palmitoylation steps are reversible. Degradation can occur from all states. (B-E): HeLa cells were transfected or not for 24h with calnexin-WT-HA, HA-calnexin-KDEL, calnexin-CA-HA, calnexin-AC-HA, calnexin-AA-HA and with or without DHHC6-myc, or additional transfected for 72h with DHHC6 siRNA. Cells were incubated 20 min pulse at 37°C with ³⁵S-methionine/cysteine, washed and further incubated for different times at 37°C. Calnexin was immunoprecipitated and analyzed by SDS-PAGE. Autoradiography (B) and western blotting were quantified using the Typhoon Imager (Image QuantTool, GE healthcare). (C) Decay profile of WT calnexin (Calx-WT) and of a calnexin mutant in which the transmembrane and cytosolic domain were removed and replaced by the KDEL sequence for ER retention (Calx-KDEL). Errors correspond to standard deviations (n=7 for Calx-WT, n=3 for Calx-KDEL). (D) Decay profile of WT calnexin was observed under normal condition (Ctrl) and after overexpression (DHHC6 Overexpression) or silencing (DHHC6 Silencing) of DHHC6. Errors correspond to standard deviations (n=7, 3, 3 for Ctrl, DHHC6 overexpression and silencing respectively). (E) Decay profile of WT calnexin (Calx-WT) and different mutants, in which site c1 (Calx-AC), site c2 (Calx-CA), or both palmitoylation sites (Calx-AA) were mutated. Errors correspond to standard deviations (n=8, 3, 3, 3 for Calx-WT, AC, CA and AA respectively).

Calnexin is synthesized by ER-associated ribosomes and inserted into the ER membrane (represented by rCAL). Presumably already co-translationally, the calnexin luminal domain undergoes folding, a process that is very efficient as discussed below. The cytosolic tail of calnexin is predicted to be highly disordered (IUPRED [97], <http://iupred.enzim.hu/>). fCAL represents folded but non-

palmitoylated calnexin. This species can be modified on the first or second palmitoylation site leading to c1CAL and c2CAL, respectively. Single palmitoylated species can acquire a second palmitate both leading to c12CAL (Figure 4.1A). The palmitoyltransferase DHHC6 catalyses the palmitoylation reaction for each site [34]. Depalmitoylation, which requires acyl protein thioesterases [47], the identity of which remain to be established for calnexin, can occur from both sites.

The individual palmitoylation and depalmitoylation steps were assumed to follow irreversible Michaelis-Menten kinetics, similar to the kinetics used in signalling network models [98]. These kinetics were determined using the total quasi steady state approximation (tQSSA) to a mass action model [75,78,99] reported by Pedersen et al. for the study of phosphorylation and dephosphorylation cycles [75,76]. This approach included a step-by-step application of tQSSA to a variety of biochemical reactions and contained a formulation of tQSSA for systems with competitive inhibition. As Gunawardena and colleagues pointed out [100] this avoids any ad hoc assumption inherent in the common Michaelis-Menten equations and takes into account sequestration effects when enzymes have multiple substrates. In these systems, a single enzyme can catalyse the same reaction on multiple substrates, which is conceptually similar to multiple sites on the same substrate. The equations could therefore be adapted to palmitoylation and depalmitoylation of the two cytoplasmic calnexin cysteines.

Although this approximation is valid for a very wide range of enzyme-substrate concentrations [76], in the case of calnexin palmitoylation, the validity of the Michaelis-Menten equations was ensured by the relative concentration of DHHC6 with respect to calnexin. Based on quantitative proteomic studies, the ratio of DHHC6-to-calnexin is indeed in the order of 1:500 [73,74,101]. Since the two sites might not be equivalent and since rates might depend on the occupancy of the neighbouring site, we introduced different K_m values for each palmitoylation and depalmitoylation step. Also since the same enzyme catalyses the palmitoylation of both sites, a competition term between the two sites was implemented in the

enzymatic kinetics as described in [75]. A similar competition term was introduced for depalmitoylation. The model also includes degradation rates for each species, with different first-order rate constants. The description of the rate expressions, the definition of the parameters, and the assumptions used in the development of the model are described in detail in the Expanded View (Figure S4.8 and Table S4.2-S4.3).

4.4.2 Experimental analysis of the effect of palmitoylation on calnexin stability

Sets of experimental data were generated to calibrate and test the predictability of the model. Since degradation of all species was included in the model, we first monitored protein degradation kinetics using ^{35}S Cys/Met metabolic pulse-chase experiments. This requires immuno-precipitation of calnexin which we performed either using a polyclonal antibody against the C-terminus to follow the endogenous protein, or using an anti-HA antibody to follow transiently transfected WT and calnexin-HA mutants. WT calnexin, endogenous or HA-tagged, showed identical biphasic decays, with a $t_{1/2}$ of ca. 8h (shown only for calnexin-HA Figure 4.1B-E).

The faster initial decay made us wonder what the contribution is of the luminal domain, the largest domain of calnexin, in shaping this decay curve. We generated a truncated version of calnexin consisting of the luminal domain fused to the KDEL sequence to ensure ER localization. HA-calnexin-KDEL was far more stable than the full-length protein (Figure 4.1BC). In particular there was no decay at early time points indicating that the luminal domain undergoes efficient folding. Thus the initial decay phase observed for the full-length protein must be dictated by transmembrane domain and/or the cytosolic tail of calnexin.

Decay of full-length calnexin was accelerated by silencing the DHHC6 enzyme (Figure 4.1D), or by mutating one or both of the palmitoylation sites (Figure 4.1E). In reverse, calnexin decay was slowed down by overexpression of DHHC6 (Figure 4.1D). Based

on these metabolic 20 min pulse-chase experiments, the apparent half-life of WT calnexin is ca. 8h, while that of non-palmitoylated calnexin, obtained either through enzyme silencing or site mutation, is ca. 5h.

4.4.3 Experimental determination of palmitate turnover on Calnexin

To further fuel the mathematical model of the palmitoylation cycle with experimental data, we determined the turnover of the palmitate moiety once attached to calnexin. To do so, we performed pulse-chase experiments using ^3H -palmitate, followed by immunoprecipitation of calnexin. Following a 2 h ^3H -palmitate pulse, loss of ^3H -palmitate from WT calnexin occurred with an apparent half-life of ca. 8 hrs and was somewhat accelerated in single cysteine mutants (Figure 4.2AB).

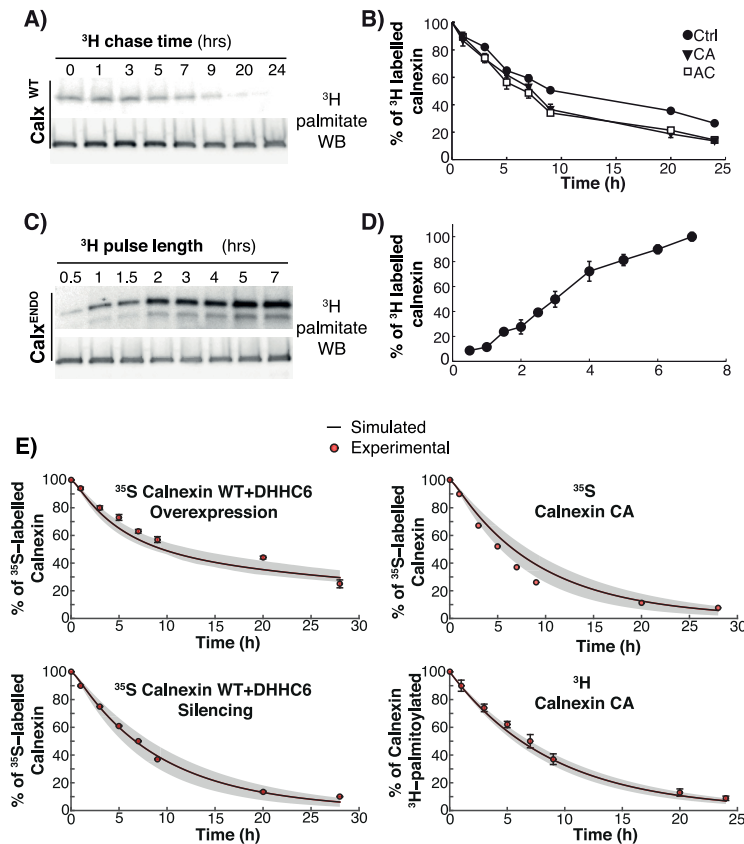


Figure 4.2. Experiments and modelling of Palmitoylation/depalmitoylation of calnexin. (AB) HeLa cells were transfected for 24h with calnexin-WT-HA, calnexin-CA-HA, or calnexin-AC-HA. Cells were incubated with ^3H -palmitic acid for 2h, washed and further incubated for different times at 37°C in complete medium prior to immunoprecipitation using anti-calnexin or anti-HA antibodies. Immunoprecipitates were split into two, run on SDS-PAGE and analyzed either by autoradiography (^3H -palmitate) or Western blotting (anti-HA). Autoradiograms were quantified using the Typhoon Imager (Image QuantTool, GE healthcare). Errors correspond to standard deviations ($n=3$). (CD) HeLa cells were incubated with ^3H -palmitic acid for different times at 37°C , washed prior to immunoprecipitation using anti-calnexin antibodies. Immunoprecipitates were split into two, run on SDS-PAGE and analyzed either by autoradiography (^3H -palmitate) or Western blotting (anti-calnexin). Errors correspond to standard deviations ($n=4$). (E) Validation of the model output through comparison of the *in silico* experiments with experimental data that was not used in the calibration of the model. Solid line is the mean of the simulations of 382 models; the shaded area is defined by the first and the third quartile of the simulations of 382 models (details of the *in silico* labelling experiment can be found in the Supplemental Information).

We also monitored the kinetics of incorporation of ^3H -palmitate into calnexin (Figure 4.2CD). This was performed in the presence of cycloheximide, to prevent synthesis of new proteins during the labelling time. Degradation was not prevented with any drug. Palmitate incorporation increased as a function of time and did not reach a plateau within the 7 h time frame of the experiment. Experiments were kept within this time frame to avoid toxic/indirect effects of prolonged inhibition of protein synthesis.

Altogether these experiments show that WT calnexin can undergo palmitoylation hours after it has been synthesized and that turn over of palmitate is very slow.

4.4.4 Calibration and validation of the mathematical model

A subset of the above ^{35}S Cys/meth and ^3H -palmitate pulse-chase experiments were used to calibrate the system, i.e. for parameter estimation, namely: decays of WT, the AC and the double AA mutants, incorporation of ^3H -palmitate into WT calnexin, ^3H -palmitate turnover for WT and the CA mutant. Since there is not a unique set of parameters that fits such a dataset, we employed a stochastic optimization method, which allowed us to generate a population of models, each having different combinations of parameter values while all being consistent with the calibration experiments. From a population of 10'000 models, we selected 382 models that best fitted the experimental data used as objective function (see Expanded View, results of the fitting on the calibration dataset are shown in Figure S4.8A-F). The pool of selected models was subsequently used for the simulations and analyses. Note that all generated predictions were obtained by simulating each model independently. The outputs of all the models were then averaged and the standard deviation with respect to the mean was used as a measure of the variability among the different models (Figure S4.8A-L).

The remaining set of experiments were used to validate the output of the model, i.e. test its predictability (see Expanded View), namely the WT calnexin decay upon DHHC6 silencing or overexpression, the decay of the single CA cysteine mutant as well as palmitate turnover for this mutant. As illustrated in Figure 4.2E and Figure S4.8G-L, the predictions were in close agreement with the experimental data, indicating that the model reliably describes the events of the calnexin palmitoylation/depalmitoylation cycle.

4.4.5 Predicted species distribution of calnexin

The model was first used to determine the distribution, and the evolution thereof, of the 5 palmitoylation species during the ^{35}S -Cys/Meth pulse-chase experiments. We set up an *in-silico* 20 minutes labelling experiment (see Figure S4.9) and calculated the relative concentrations of calnexin in the different palmitoylation states at different time points of the chase (Figure 4.3A).

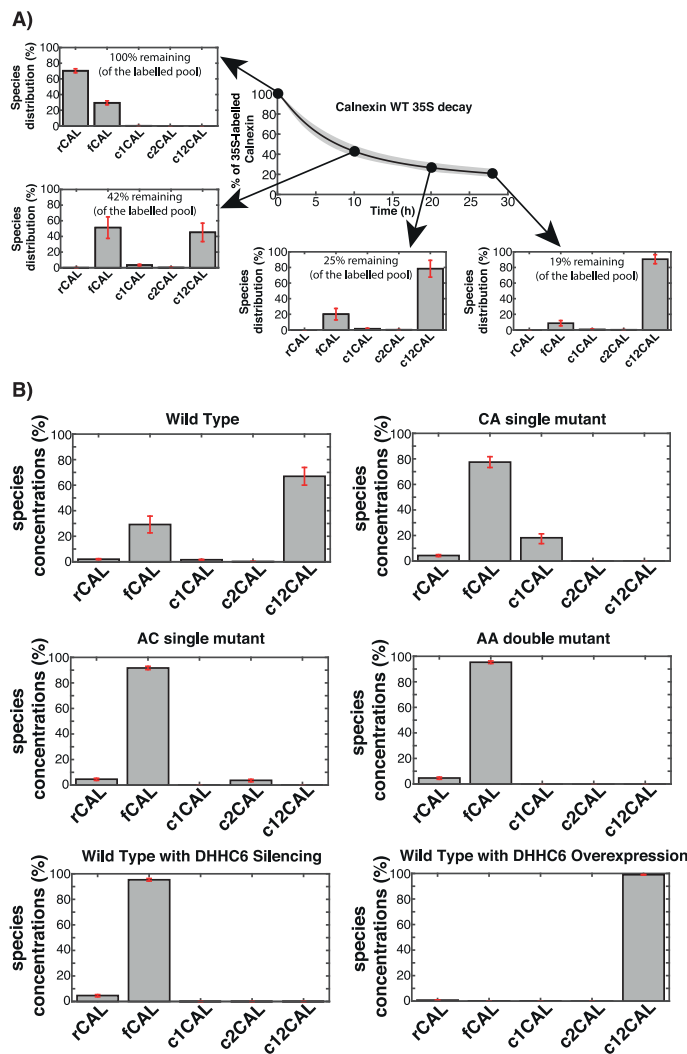


Figure 4.3. *In silico* analysis of the calnexin species distribution. (A) We replicated *in silico* the ^{35}S pulse-chase experiment on WT calnexin. During the experiment we monitored the relative distribution of calnexin in the different palmitoylation states. Solid line is the mean of the simulations of 382 models; error bars are defined by the first and the third quartile of the simulations of 382 models (details of the *in silico* labelling experiment can be found in the Expanded View). (B) The model was used to predict the steady state distribution of WT calnexin and of the mutants in

the cell. rCAL correspond to calnexin during the synthesis, fCAL represent folded calnexin while c1CAL and c2CAL denote the two single palmitoylated species. c12CAL represents the double palmitoylated calnexin. Error bars correspond to first and third quartile of the simulations of 382 models (details of the *in silico* labelling experiment can be found in the Supplementary Information Information).

At the end of the metabolic pulse ($t=0$), the population is predicted to be exclusively of non-palmitoylated, ca. 25% of which are already folded. Ten hours after the pulse, the entire calnexin population was predicted to be folded, as can be expected, but, more unexpectedly, ca. 50% was still non-palmitoylated (Figure 4.3A). Almost the palmitoylated species, the dually-palmitoylated was predicted to be the most populated, with barely any single palmitoylated species. As chase time proceeded, the non-palmitoylated species decreased again to the benefit of the dual palmitoylated form, which approached 90% at the end of the chase period (Figure 4.3A). Note that distributions are expressed as percentages of the remaining population, not of the initial population, which decreased by 80% between the beginning and the end of the chase period.

This analysis indicates that the initial faster phase of degradation occurs when most of calnexin is non-palmitoylated. Once a significant percentage of the population becomes acylated, the rate of decay drastically decreases, confirming a stabilizing effect of palmitoylation.

We next determined the distribution of WT calnexin at steady state, since it may differ from the distribution at the end of our chase period. At steady state in our HeLa cells and under our experimental conditions, the model predicts that ca. 70% of calnexin is dual-palmitoylated and the remaining population is free of palmitate. The prediction that the great majority of cellular calnexin is palmitoylated is consistent with our previous experimental evidence [34].

We next utilised the model to estimate the steady state distribution of calnexin species in cells over-expressing DHHC6 and found that the dually-modified population increased to almost 100% while it fell to zero upon silencing of the enzyme (Figure 4.3B), as expected. By regulating DHHC6 amounts and/or activity, cells thus have the potential to control the percentage of calnexin that is dually-palmitoylated, and thus functional.

Consistent with the low population of single palmitoylated species for WT calnexin, determination of the steady state species distribution of the CA and the AC calnexin mutants revealed that >80% of the molecules are non-palmitoylated. This indicates that to obtain a significant population of palmitoylated calnexin, two sites are necessary. That the single palmitoylated species do not get significantly populated was initially unexpected given the relatively small difference in the apparent palmitate turnover rates of WT and single cysteine mutants (Figure 4.2B). To understand this apparent inconsistency, we used the model to predict the distribution of the species labelled during the 2 hrs ^3H -palmitate pulse (corresponding to $t=0$ in Figure 4.2B) (Figure S4.10, Figure S4.11, Supplementary material). *In silico*, ^3H -palmitate labelling shows that in 2 hrs, only a minute percentage of the total WT population –less than 3%– undergoes labelling (Figure S4.11). This is consistent with the fact that 70% of the steady state population is dual palmitoylated at steady state and can thus not be further modified upon addition of ^3H -palmitate. Moreover, Figure 4.2D indicates that palmitoylation can be very slow, since some proteins undergo palmitoylation 7 h or more after having been synthesized (Figure 4.2D). Thus Figure 4.2B represents the loss of ^3H -palmitate from a population of ^3H -palmitate labelled species composed roughly of 50% c1CAL and 50% c12CAL (Figure 4.4A), jointly representing just 3% of the total population.

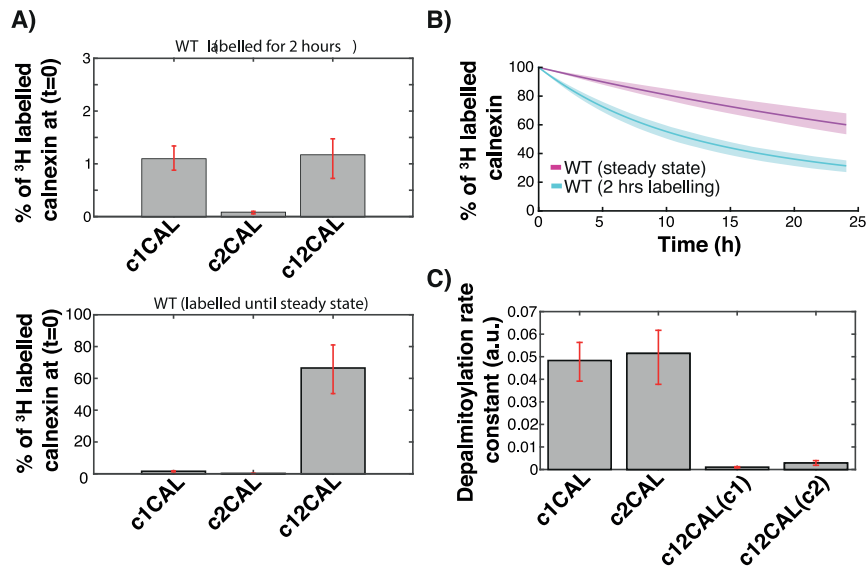


Figure 4.4. Prediction of depalmitoylation kinetics. (A) WT calnexin was labelled with ³H-palmitate *in silico* either for two hrs or to reach the steady state distribution of palmitoylation species. (B) The kinetics of palmitate loss starting from the two distributions in A were determined. The solid lines represent the means of the simulations of 382 models and the shaded areas are defined by the first and the third quartile of the simulations of 382 models (details of the *in-silico* labelling experiment can be found in the Expanded View). (C) The rate constants for depalmitoylation from c1CAL, c2CAL and c12CAL were determined. Error bars correspond to first and third quartile of the simulations of 382 models

To estimate the rate of palmitate release from dually-palmitoylated calnexin (c12CAL), which is not readily accessible experimentally, we determined what the ³H-palmitate decay would be if the starting point was that of steady state palmitoylation, i.e. 70% calnexin dually labelled (Figure 4.4A, bottom panel). Under these conditions, the loss of palmitate was significantly slower (Figure 4.4B). In fact, we measured an average turnover rate for palmitate of ~32h, which is almost 3 times slower than the apparent turnover rate estimated by our 2 hrs labelling experiment. This analysis indicates that the rate constant of palmitate removal from a given site depends on the occupancy of the second site. Our predictions indeed indicate that, in situations of single site occupancy, the rate constants of loss of palmitate from site 1 are drastically higher than the rate constants of loss from either site in the situation of double occupancy (Figure 4.4C). Thus, single palmitoylated species do not get significantly populated as compared to the double palmitoylated one, because they lose their palmitate at a far higher rate.

Altogether this analysis indicates that two sites are required to stably palmitoylated calnexin, because double occupancy drastically slows down depalmitoylation.

4.4.6 Estimated half-lives of palmitoylated calnexin species

³⁵S Cys/Met pulse-chase experiments of WT and cysteine mutants indicate that palmitoylation has a stabilizing effect and that the half-life of palmitoylation deficient calnexin is ca. 5 hrs. They however do not allow direct identification of the half-life of single or dually palmitoylated calnexin. We therefore made use of the model to predict the decay kinetics of each of the calnexin species (Table 4.1 and Figure 4.5A).

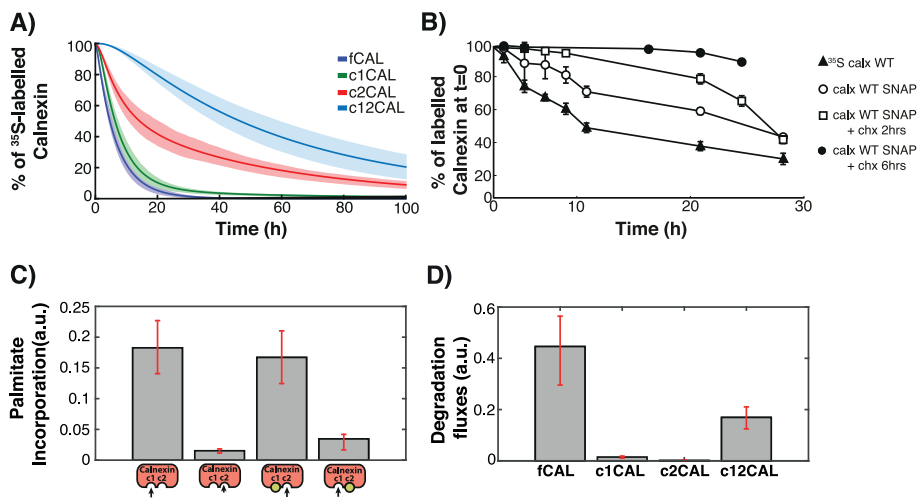


Figure 4.5. Kinetics of calnexin palmitoylation and stability. (A) The decay curves were predicted for the different palmitoylation species: fCAL represent folded calnexin, c1CAL and c2CAL the single palmitoylated species, c12CAL double palmitoylated calnexin. Solid line is the mean of the simulations of 382 models; shaded area is defined by the first and the third quartile of the simulations of 382 models (details of the *in silico* labelling experiment can be found in the Expanded View). (B) Decay profiles were determined for WT calnexin either by ³⁵S Cys/Meth labelling (³⁵S calx WT) and by SNAP labeling (Calx WT SNAP). (C) The model was simulated until it reached steady state and the values of the palmitoylation rates under steady state conditions were plotted. Error bars correspond to first and third quartile of the simulations of 382 models (details of the *in silico* labelling experiment can be found in the Expanded View). (D) The model was simulated until it reached steady state and the values of the degradation rates under steady state conditions were plotted. Error bars were determined as in (C).

Palmitoylation of site 1 has a mildly stabilizing effect, the half-life of c1CAL being ca. 6.5 hrs (Figure 4.5A). Palmitoylation of site 2, whether site 1 is or not modified, leads to a spectacular stabilization, with dual-palmitoylated calnexin having a predicted half-life of 45 hrs (Table 4.1 and Figure 4.5A).

We sought for an experimental confirmation for this almost 10-fold increase in protein stability. At steady state, 70% of calnexin is predicted to dually-modified (Figure 4.3B). To assess the half-life of this population, we chose to monitor the stability of calnexin tagged with SNAP, a widely used protein that self-labels when incubated with O^6 -benzylguanine derivatives [102]. We labelled cells for 30 min with SNAP-cell-TMR-star, a red fluorescent substrate of SNAP, in order to label the entire, steady state, population of calnexin-SNAP. Following different periods of chase, cells were harvested and the level of fluorescent calnexin was monitored by SDS-PAGE and fluorescence scanning. As shown in Figure 4.5B, the decay of calnexin-SNAP-TMR-star was slower than that observed by ^{35}S pulse-chase, yet still biphasic. At $t=0$ of SNAP-cell-TMR-star labeling, the population of calnexin-SNAP molecules have different "ages", the "youngest" having just been synthesized. We therefore tested whether pretreatment of cells with cycloheximide for different times, to block protein synthesis, would affect the apparent stability of the SNAP-cell-TMR-star labeled population. A 2 hrs cycloheximide pretreatment already led to an increase in apparent half-life of the population (Figure 4.5B). We extended the pretreatment to 6 h, so that all calnexin molecules would be at least 6 h "old". With this treatment, no decay was observed for the first 24 hrs, and this was followed by a slow decline, leading to an apparent half-life of ca. 47 hrs (Figure 4.5B).

Altogether, the mathematical modeling and the calnexin-SNAP tagged decay analysis indicate that palmitoylation of the two juxtamembranous sites leads to a dramatic increase in the stability of calnexin.

Table 4.1. Half-life of the different calnexin species. The standard deviation is estimated from the 382 models (see Supplementary material).

Species	Half-Life
C^{00}	5.0 ± 1.5
C^{10}	6.5 ± 1.8
C^{01}	14.5 ± 1.8
C^{11}	46.0 ± 11.2

4.4.7 Cooperativity between the two calnexin palmitoylation sites

The analysis of the steady state distribution of WT and cysteine mutants (Figure 4.3B) indicates that two sites are required for stable palmitoylation and suggest that there might be cooperativity between sites. We therefore estimated the rates of palmitate incorporation at each site, depending on whether the other site was occupied or not. As shown in Figure 4.5C, palmitoylation on site 1 is predicted to be drastically more efficient than on site 2. However, if site 1 is already occupied, site 2 is readily modified. In marked contrast, if site 2 is occupied, site 1 by being positioned between the transmembrane and the palmitoylated site 2 might be less accessible to the enzyme. This reaction flux analysis indicates that site 1 is preferentially modified and when this has occurred, site 2 is rapidly acylated, indicating positive cooperatively between sites 1 and 2.

4.4.8 Efficient degradation of calnexin is predicted to require depalmitoylation

We next simulated the degradation fluxes of the different calnexin species (Figure 4.5D). Even though nearly 70% of the protein is dual-palmitoylated at steady state (Figure 4.3B), the degradation flux was highest for the non-palmitoylated state. Double-palmitoylated calnexin did undergo degradation, but at a 3 to 4 times slower rate, which is due to the 3 to 4 times lower values of the degradation rate constants (by parameter estimation using the corresponding experimental information, Table S4.3). Therefore, efficient degradation of calnexin appears to require prior depalmitoylation by thioesterases.

4.4.9 Estimation of the lag-time separating calnexin synthesis from palmitoylation

An unexpected and intriguing observation in this study is the slow appearance of the palmitoylated population following synthesis (Figure 4.3A) and the absence of a plateau in the palmitate incorporation experiment (Figure 4.2). This suggests that there is a lag time between synthesis and palmitoylation of calnexin. To evaluate this lag time, we derived a stochastic formulation of the original model of the palmitoylation process, and performed stochastic simulations that allowed us to track single proteins in the system from synthesis to dual-palmitoylation (Supporting Material, Figure S4.12). Simulations were performed for 5000 molecules. Of these about 3000 were degraded before any palmitoylation event occurred. For the remaining population, we determined the frequency distribution of the time required to reach dual-acylation (Figure 4.6). Some molecules underwent palmitoylation within a few hours of synthesis (Figure 4.6). Most however remained in the non-

palmitoylated form for extended periods of time, leading to an average time of synthesis-to-palmitoylation of 8 hrs (Figure 4.6).

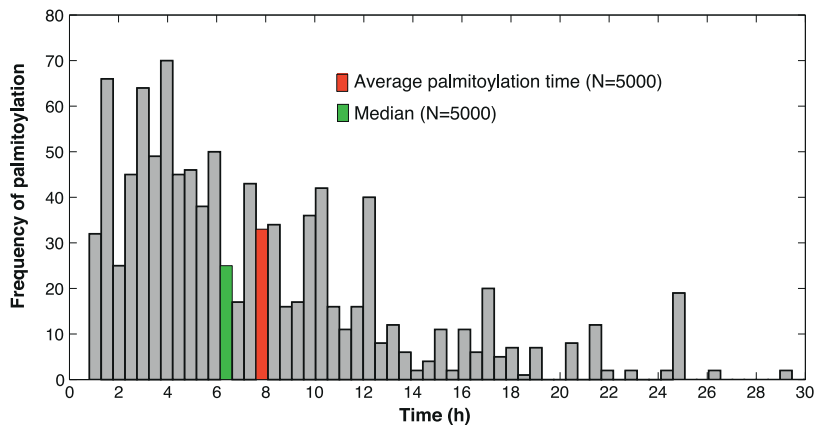


Figure 4.6. Stochastic simulations reveal the average palmitoylation time for calnexin. Single proteins were tracked in 5000 stochastic simulations of the labelling method described in Expanded View . From the simulations we estimated the average and median time requested for a single molecule of calnexin to undergo double palmitoylation.

4.4.10 Phosphorylation is a negative regulator of calnexin palmitoylation

Calnexin and its palmitoylating enzyme DHHC6 are two membrane proteins that reside within the same two-dimensional space of the ER membrane. The lag time between synthesis and palmitoylation of calnexin could potentially be due to slow diffusion of one of the two molecules. We therefore determined their mobility using Fluorescence Recovery after Photobleaching (FRAP) of C-terminally GFP-tagged variants. Both calnexin and DHHC6 showed rapid diffusion with rates of 0.63 ± 0.09 and $0.67 \pm 0.13 \mu\text{m}^2/\text{s}$ respectively, in close agreement with previously published rates for calnexin (Figure 4.7A) [103]. As a control of a slow diffusing membrane protein, we confirmed that Climp63/CKAP4 has a diffusion rate of $0.06 \pm 0.01 \mu\text{m}^2/\text{s}$ (Figure 4.7A) [104].

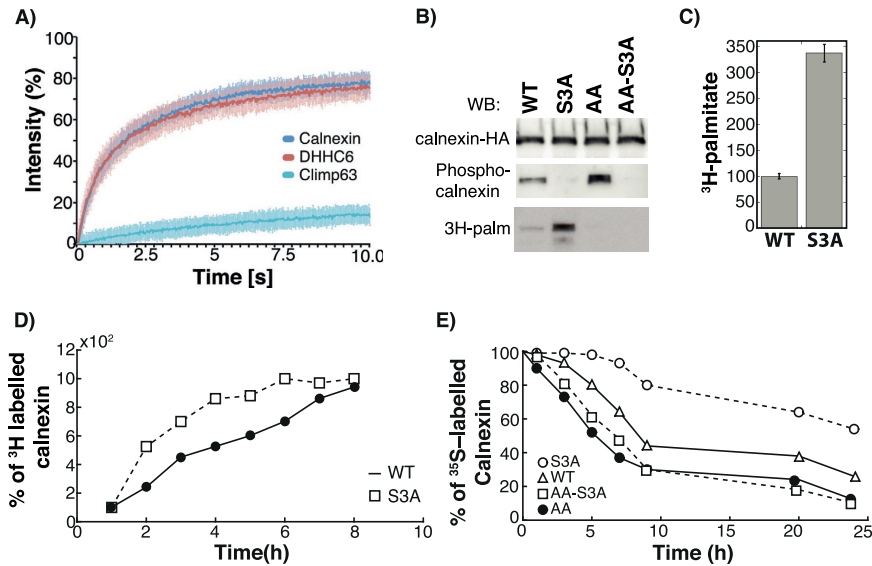


Figure 4.7. Premature calnexin palmitoylation is prevented by serine phosphorylation. (A) HeLa cells were transfected 48 hrs with Calnexin-GFP, GFP-DHHC6 or GFP-Climp63. Cells were submitted to FRAP analysis (see Material and Methods). (B-E) HeLa cells were transfected for 24h with calnexin-WT-HA, calnexin-S3A-HA, calnexin-AA-HA or calnexin-AA-S3A-HA. (C) Cells were incubated with ^3H -palmitic acid for 2h, washed prior to immunoprecipitation using anti-HA antibodies. Immunoprecipitates were split into two, run on SDS-PAGE and analyzed either by autoradiography (^3H -palmitate) or Western blotting (anti-HA or anti phospho-calnexin). Autoradiograms were quantified using the Typhoon Imager (Image QuantTool, GE healthcare). Errors correspond to standard deviations ($n=3$). (D) Cells were incubated with ^3H -palmitic acid for different hours at 37°C , washed prior to immunoprecipitation using anti-HA antibodies. Immunoprecipitates were split into two, run on SDS-PAGE and analyzed either by autoradiography (^3H -palmitate) or Western blotting and autoradiograms were quantified using the Typhoon Imager (Image QuantTool, GE healthcare). (E) Cells were incubated 20 min pulse at 37°C with ^{35}S -methionine/cysteine, washed and further incubated for different times at 37°C . Calnexin were immunoprecipitated and analyzed by SDS-PAGE. Autoradiography and western blotting were quantified using the Typhoon Imager (Image QuantTool, GE healthcare).

Given the high rates of DHHC6 and calnexin diffusion, slow palmitoylation of calnexin cannot be explained by a low probability of encounter between the enzyme and the substrate. This raised the possibility that calnexin palmitoylation is actively prevented and led us to investigate the effect of calnexin phosphorylation on its acylation. We generated a triple mutant in which all three serine phosphorylation sites in the calnexin tail (Ser- 554, 564, 583) were mutated to alanine (S3A mutant). Remarkably, mutation of the serine phosphorylation sites led to a ~200% increase in palmitoylation during the 2h labelling period (Figure 4.7BC). Also, kinetics of ^3H -palmitate incorporation, in the absence of protein synthesis, were faster for the S3A mutant than for WT calnexin (Figure 4.7D). Consistently, ^{35}S Cys/Meth pulse-chase

experiments revealed that the S3A mutant was far more stable than the WT protein (Figure 4.7E), in particular due to the disappearance of the initial rapid decay phase, which in WT is due to degradation of the non-palmitoylated species. Altogether these observations indicate that palmitoylation of calnexin is under the negative control of serine phosphorylation.

4.5 Discussion

S-palmitoylation is a post-translational modification that is receiving increasing attention as more and more key cellular events and pathways appear to rely on the reversible acylation of specific proteins [15,16]. The dynamics of this modification and the regulatory mechanisms are however poorly understood. We have here investigated palmitoylation of the key ER chaperone calnexin [88]. Calnexin shares its ability to promote folding of N-glycosylated proteins with calreticulin, a soluble ER protein. Calnexin in contrast spans the ER membrane and harbors a 90 residue cytosolic tail. While the transmembrane nature of calnexin was initially thought to favor the chaperoning of transmembrane proteins, it is increasingly clear that the transmembrane domain and cytosolic tail confer additional properties and functions to calnexin which involve its palmitoylation [89,90,92,105].

We developed a mathematical model of the palmitoylation cycle that accurately captures the properties of the system as shown by its predictive power. This combination of modeling and experimentation led to a variety of interesting, often unexpected, analyses, predictions and conclusions. First, it highlights the under-appreciated complexity of classical metabolic pulse-chase experiments. What is referred to as WT –of calnexin or any other protein– is in fact a complex population of species, the distribution of which evolves with time. In our model, we have distributed calnexin into 5 species, defined by the folding and palmitoylation status. In future studies we will increase the complexity of the model, including additional

species with different types of post-translational modifications, such as phosphorylation on three serines of its cytosolic tail [92], a modification which we find affects its palmitoylation status.

Our analysis also indicates that, while ^{35}S Cys/Met labels a well-defined sub-population –the one that has been synthesized by the cell during the pulse–, this is by no means the case when labeling cells with ^3H -palmitate. Radiolabeled palmitate can indeed be incorporated into any molecule that has an unoccupied palmitoylation site but importantly cannot be incorporated into fully palmitoylated proteins, which in the case of calnexin compose ca. 70% of the total cellular population at steady state. These fully palmitoylated proteins are thus silent in such a ^3H -palmitate pulse-chase analysis.

Most importantly our model provides unprecedented understanding of the palmitoylation process. In the context of calnexin, we found that the molecule undergoes palmitoylation first on site 1. Once this has occurred, site 2 can be readily modified. There is thus cooperativity between site 1 and 2. Depalmitoylation can rapidly occur if a single site is occupied but removal is drastically slowed down if the two sites are acylated. As a consequence, calnexin is either not modified or acylated on both sites, the single palmitoylated state being barely populated. Importantly, the percentage of calnexin that is modified at steady state can be tuned between 0 to 100% by the activity of the DHHC6 palmitoyltransferase. It has recently been shown that the activity of DHHC6, at least for certain substrates, depends on its association with Selenoprotein K (selK) [106]. Selk was found to be upregulated by ER stress [107], i.e. when increased chaperone activity is necessary, which would lead to an increase in DHHC6 activity and thereby in calnexin protein.

Interestingly, we found that palmitoylation not only affects calnexin function – through association with the ribosome-translocon complex or ER-mitochondrial interaction sites [58,89,90,92]– but has a drastic effect on its stability, increasing the average half-life of the molecules from 5 to 45hrs. Interestingly however, palmitoylation only occurs, on average, 8 hrs after the calnexin molecule has been

synthesized. This means that, at the population levels, some 80 % of the calnexin molecules that a resting cell synthesizes are degraded before they acquire palmitate and thereby activity. Considering that a resting cell has 500'000 to 1 million copies of calnexin at steady state [73,74,101], this implies that 2.5 to 5 millions copies were actually synthesized. While 80 % of these are degraded in resting cells, they can potentially be mobilized if palmitoylation occurred earlier after synthesis.

This unexpected apparent inefficiency for a protein that has no folding problems (Figure 4.1C), combined with the fact that both calnexin –the substrate– and DHHC6 –the enzyme– rapidly diffuse within the same membrane and thus must have a high probability of encounter, led us to search for a mechanisms that would actively prevent palmitoylation.

Since calnexin is known to undergo serine phosphorylation, we investigated the potential impact of this modification on palmitoylation and found that a serine phosphorylation deficient calnexin mutant undergoes greatly accelerated palmitoylation. Our study thus indicates that cells can post-translationally tune the expression level of calnexin by controlling the kinetics of its palmitoylation via phosphorylation of its cytosolic domain.

In addition to providing important novel insight into the mechanisms by which cells control the level of chaperone activity in the ER, the mathematical model that we have elaborated provides a framework to study palmitoylation of other proteins. In particular, a great variety of type I membrane proteins harbor one or two palmitoylation site in the vicinity of their transmembrane domains. It will be of interest to determine whether the rules of cooperativity and 2 site-requirement revealed by calnexin apply to other proteins and how palmitoylation kinetics of these are controlled. Considering that palmitoylation and depalmitoylation are mediated by enzymes that maybe themselves undergo cycles of palmitoylation and depalmitoylation [108,109], this post-translational modification is particularly in need of mathematical modeling to understand the complexity of its regulation.

4.6 Material and Methods

4.6.1.1 Cells, Antibodies and Reagents

Hela cells were grown in complete MEM (Sigma) supplemented with 10% foetal bovine serum (FBS), 2 mM L-glutamine, penicillin and streptomycin. Rabbit antibodies against calnexin were produced in our laboratory against the C-terminal peptide: CDAEEDGGTVSQEEEDRKPK; anti phospho S583-calnexin were from Abcam (ab58503), anti-HA and anti-HA-agarose conjugated beads were from Roche (Applied Science, IN), protein G-agarose conjugated beads from GE Healthcare, HRP secondary antibodies from Pierce.

4.6.1.2 Plasmids and transfection

Human calnexin-HA and human calnexin-C501A-C502A-HA, calnexin-C501A-HA (AC), calnexin-C502A-HA (CA), calnexin S554A-S564A-S583A-HA (S3A), calnexin-C501A-C502A-S554A-S564A-S583A-HA (AAS3A) and human-calnexin-SNAP-HA, dog-HA-calnexin-KDEL and human DHHC6-myc were cloned in pcDNA3 [108,109]. For control transfections, we used an empty pcDNA3 plasmid. Plasmids were transfected into Hela cells for 24 hrs (2 µg cDNA/9.6 cm² plate) using Fugene (Roche Diagnostics Corporation). To generate GFP fusions, calnexin and DHHC6 were cloned into the pGFP vector and CKAP4 was cloned into the peCE vector.

4.6.1.3 RNAi experiments

siRNA against human DHHC6 were purchased from Qiagen (target sequence: gaggtttacgatactggttat). As control siRNA we used the following target sequence of the viral glycoprotein VSV-G: attgaacaaacgaacaagga. For gene silencing, Hela cells were transfected for 72 h with 100pmol/9.2cm² dish of siRNA using interferin (Polyplus) transfection reagent.

4.6.1.4 Radiolabeling experiments

To follow palmitoylation, calnexin expressing cells were incubated 2h or several hours at 37°C in incubation medium (Glasgow minimal essential medium buffered with 10 mM Hepes, pH 7.4) with 200 μ Ci /ml 3 H palmitic acid (American Radiolabeled Chemicals, Inc), washed and incubated different times at 37°C with complete medium prior to immunoprecipitation using anti-calnexin or anti-HA antibodies. Beads were incubated 5 min at 90°C in reducing sample buffer prior to SDS-PAGE. Immunoprecipitates were split into two, run on 4-20% gels and analyzed either by autoradiography (3 H-palmitate) after fixation (25% isopropanol, 65% H₂O, 10% acetic acid), gels were incubated 30 min in enhancer Amplify NAMPI00 (Amersham) and dried or submitted to Western blotting (anti-calnexin). Autoradiograms and western blotting were quantified using the Typhoon Imager (Image QuantTool, GE healthcare). For metabolic labeling, Hela cells were transiently transfected (24h) or not with calnexin-HA cDNAs, washed with methionine /cysteine free medium, incubated 20 min pulse at 37°C with 50 μ Ci/ml 35 S-methionine/cysteine (Hartman Analytics), washed and further incubated for different times at 37°C in complete medium with a 10-fold excess of non-radioactive methionine and cysteine. Calnexin were immunoprecipitated and analyzed by SDS-PAGE. Protein synthesis was blocked by 30 min treatment with 10 μ g/ml cycloheximide (Sigma) at 37°C.

4.6.1.5 Immunoprecipitation

For immunoprecipitations, cells were lysed 30 min at 4°C in IP buffer (0.5%NP40, 500 mM Tris-HCl pH 7.4, 20 mM EDTA, 10 mM NaF, 2 mM benzamidine, and a cocktail of protease inhibitors, Roche), centrifuged 3 min at 2000 g and supernatants were pre-cleared with protein G-agarose conjugated beads and supernatants were incubated 16 h at 4°C with antibodies and beads.

4.6.1.6 SNAP labeling

To follow SNAP labeling, calnexin-SNAP expressing cells were incubated 6 hours with 10 µg/ml cycloheximide at 37°C in complete medium, then 30min at 37°C in complete medium with 1 µM SNAP-cell-TMR-star (Biolabs) and cycloheximide, then washed three times and incubated different times at 37°C with complete medium prior lysis of the cells. Total cell lysates were analyzed by SDS-PAGE and the fluorescence was measured using the Typhoon Imager (Image QuantTool, GE healthcare).

4.6.1.7 Fluorescence Recovery After Photobleaching

HeLa cells were seeded on FluoroDish (glass bottom: 0.17mm thickness, from World Precision Instrument Inc. USA) and GFP-tagged proteins (calnexin, DHHC6 and Climp63) were transiently expressed for 24 hours. Fluorescence recovery after photobleaching (FRAP) experiments were performed on a Leica SP5 microscope using a 63x oil-immersion objective (1.4 NA). The microscope was operated using the software supplied with the instrument (LAS AF 2009). The 488nm line of the argon laser was set at 60% output and 100% transmission. During the experiment, the cells were kept in a chamber at 37°C and 5% CO₂. The pinhole was wide-open. The scanner speed was set at 1400Hz. The digital zoom was set at 6. The detector gain was set at 740V. The frame size was set at 512x32. The resulting scanning time was 32ms per frame. Point bleach measurements were performed and the effective radius of the bleached area was calculated as 0.9µm according to [110]. 50 iterations using 4% transmission of the laser were acquired as pre-bleach reference scans. Then the GFP-tagged proteins localized in the ER were bleached a single iteration for 5ms. Afterwards 500 post-bleach iterations were acquired with same settings as for the pre-bleach acquisitions. FRAP experiments were conducted for each condition on 12 different cells expressing the same level of GFP-tagged proteins. Longer post-bleach acquisitions were conducted for Climp63 to enable a better exponential fitting of the post-bleach curves. Diffusion coefficients were extracted (according to [111]) from the fitting of an exponential equation to the post-bleach recovery curves using a hand-

made code in matlab (MATLAB 8.0 and Statistics Toolbox 8.1, The MathWorks, Inc. USA).

4.7 Supplementary Material

4.7.1.1 Description of the model

One of the most important improvements in this work with respect to what is observed in different models describing covalent modification systems, is the introduction of synthesis and degradation terms. In our opinion these terms help better describe the steady state of a living organism with respect to a closed model [112], which will eventually reach a chemical equilibrium. In this model (Figure 4.1), we assume that calnexin is first synthesized by the ribosome from its mRNA to the corresponding peptide and inserted into the membrane (U). Unfolded calnexin goes through a process of folding, forming folded calnexin, which then shows the two sites available for palmitoylation (C^{00}). These two sites can be palmitoylated by DHHC6. The first palmitoylation can occur on both sites; C^{10} and C^{01} denote the two different palmitoylation states. After the first PTM calnexin can undergo another palmitoylation event still catalysed by DHHC6; C^{11} represents the double palmitoylated calnexin. Since palmitoylation is reversible the palmitate can be removed from the two sites. The removal of palmitate requires a different type of enzyme, namely Acyl Protein Thioesterases (APTs). The following assumption were considered during the design of the model:

1. Calnexin can be degraded in each of its palmitoylation states after folding. During the folding process (U to C^{00}) no degradation can occur.
2. Palmitoylation increases the half-life of calnexin (observed through experimental data).
3. DHHC6 is present in low concentrations with respect to its substrate calnexin [74].

4. The two palmitoylation sites may have a different affinity with respect to the palmitoylation/depalmitoylation enzymes, so we adopted separated K_m s for the two sites.
5. The catalytic constant for each palmitoylation site is the same.
6. Both palmitoylation steps are reversible. APT catalyses the depalmitoylation steps.

Despite its simplicity the model has shown to be able to describe the basic mechanisms underlying the observed experimental results (see section “Parameterization of the model”). A detailed account of all reactions and differential equations of the model is given in Table S4.2-4.

4.7.1.2 Parameterization of the model

Due to the lack of kinetic data on the enzymatic process of palmitoylation, all the parameters of the model were estimated using a genetic algorithm (GA). GA is a heuristic global optimization algorithm that mimics the process of natural selection. The detailed explanation on how parameter estimation was performed can be found in Chapter 3 - . A comparison between model predictions and experiments after optimization is visible in Figure S4.8, while the values of the estimated parameters are contained in Table S4.3.

4.7.1.3 Sensitivity analysis:

Once the calibration of the model was complete we performed parametric sensitivity analysis (SA), with the aim of gaining more insight on how the various parameters of the model affected its output. In particular, we were interested in understanding which parameters are fundamental in determining the abundance and the half-life of calnexin. SA was performed using a simple One-at-a-time (OAT) method, which consist in changing one parameter at a time and observing its effect on the states of

the model. The detailed procedure used to perform sensitivity analysis is illustrated in Chapter 3 -

Results of SA on the abundance of calnexin WT and mutants in steady state can be seen in Figure S4.13. For each mutant in the top panel we reported the steady state distribution and abundance, in the second panel we found the top 5 sensitivity indexes, while in the last panel we reported sensitivity for all the parameters. Figure S4.14-17 show sensitivity towards calnexin's half life for WT and mutants. In the top-left panel we find the measured half-life, in the top right panel we show the sensitivity indexes at 50% calnexin concentration. In the bottom left panel we find the top 5 dynamic sensitivities, while in the bottom right panel we find dynamic sensitivities for each parameter.

The top 3 sensitivity indexes for the mutants AA AC and CA show a similar pattern. In fact, the main determinant of calnexin abundance is synthesis rate *vs.* We then find the degradation rate constant of fCAL (*kdl*), followed by the folding rate of calnexin *kf*, or the degradation rate constant of C¹⁰ and C⁰¹, namely *kd2*. In these mutants palmitoylation parameters have no effect on protein abundance. While this is expected for the AA mutant, who can't be palmitoylated, it is surprising in the case of the CA and AC mutant, in which one site can still be palmitoylated. Therefore, results suggest that a single modification of calnexin have little to no effect on its abundance. Instead, if we look at the sensitivity results for the WT we can see that among the top 3 sensitivity indexes, in 2nd and 3rd position we find the degradation rate constant of c12CAL (*kd3*) and the maximum speed of palmitoylation (*Vf*). Since the concentration of WT in the steady state is higher than the one of the mutants (Figure S4.13), sensitivity analysis confirm that palmitoylation parameters are strikingly important to reach such levels of concentration. It is important to highlight that in WT the enzymatic parameters have an important role in determining its concentration, but the same parameters have almost no effect on CA and AC mutants concentration. This suggests that a single palmitoylation step is not enough to increase calnexin

concentration with respect to the AA mutant, which can't be palmitoylated. A double modification of calnexin seems to be required to reach the level of abundance of WT.

Similar results are observed when we calculate sensitivity indexes with respect to calnexin's half-life (Figure S4.14-17). Again in WT (Figure S4.14) the maximum speed of palmitoylation is among the most sensitive parameters, highlighting the important role of palmitoylation for calnexin in determining both the abundance and half-life. Looking at the sensitivities for the CA and AC mutants (Figure S4.16-17), it is easy to see that a single palmitoylation modification is not enough to affect significantly calnexin's half-life (since V_{max} and the enzymatic parameters of the model have almost no effect in CA and AC mutants, but they show an effect on WT).

4.7.1.4 Stochastic model formulation:

During the development of this model we had to deal with one main issue. Calnexin is a very abundant protein in the cell, two big proteome studies [21,40,41] have estimated that calnexin is present in HeLa cell in a number of copies between $5 \cdot 10^5$ and 10^6 . Since we wanted to measure a property of the system in steady state condition, we had to simulate our model with a number of molecules that matched the values found experimentally. The high number of molecules used, made stochastic simulations computationally expensive. Therefore, in order to perform simulations in an acceptable amount of time and resources we used the same reduction approach as we did in the deterministic model, using a model described with tQSSA. Different studies [113,114] have shown that tQSSA is an approximation that can be applied to stochastic models without losing accuracy. Validation results of the method have already been obtained for the Goldbeter switch, MAPK cascade, Michaelis-Menten kinetics and others kinetic models [114].

The model that we designed is similar to the one used for deterministic simulations, but as required by the stochastic solver, we provided a representation through a stoichiometric matrix (Table S4.5) and propensity function (Table 4.6). All

computations were performed using MATLAB R2014b. The stochastic simulations were performed using the first reaction method, a variant of the Gillespie algorithm [87].

A detailed description on model formulation, parameter conversion and setting up of the stochastic simulations is present in Chapter 3 - The stoichiometric matrix of the stochastic model, the propensity function used for simulations and the values of the converted parameters are visible in Table S4.5-7.

4.8 Appendix A

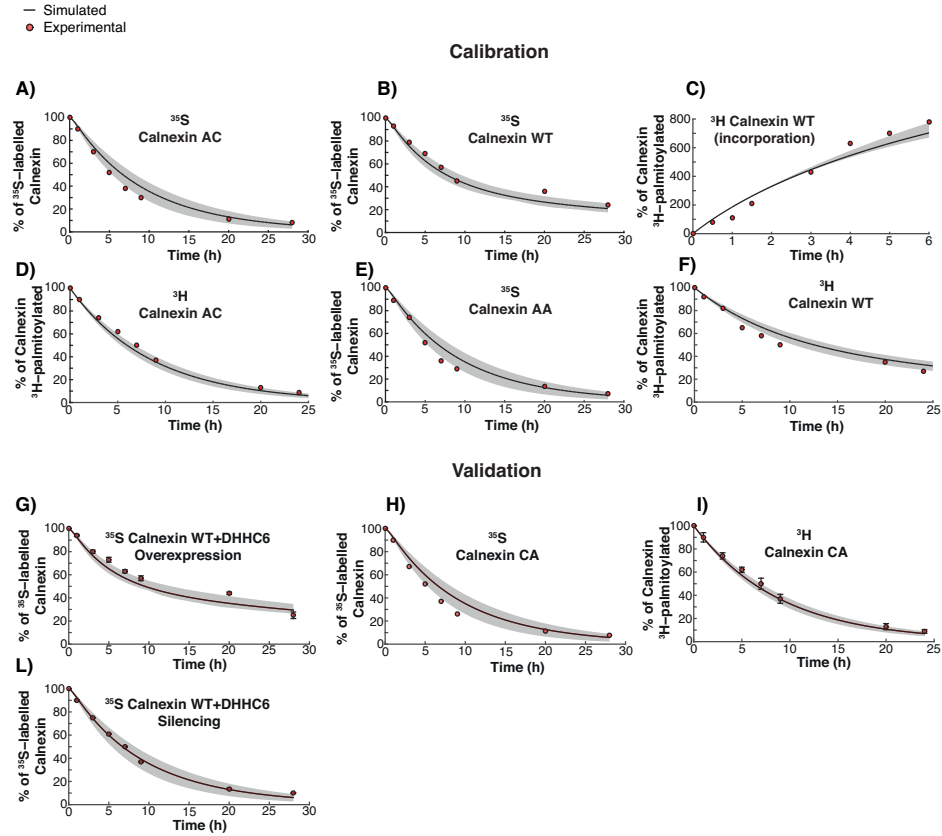


Figure S4.8. Results of the GA optimization are plotted on top of the experimental data used as objectives. (A-F): In order to estimate the parameters of the model the following set of experiments were used. The algorithm used for estimation is *gamultiobj* from MATLAB's global optimization toolbox. This algorithm finds the overall minimum of multiple objective functions by minimizing the difference between the output of the model and the experimental data of the calibration set. The goodness of the fitting is visualized plotting the output of the *in silico* experiment obtained with the model on top of the corresponding experimental data: A: AC mutant ^{35}S decay; B: WT ^{35}S decay; C: WT ^3H incorporation (data are relative to the $t=1\text{h}$ point); D: AC mutant ^3H decay; E: AA mutant ^{35}S decay; F: WT ^3H decay. (G-L): In order to ensure that the model was able to capture behaviors of experiments for which it wasn't trained we used the following set of experiments to validate the output of the model. These experiments were not used to estimate the parameters of the model, therefore its output is a prediction of the result of the different experiments. This analysis ensures that the model is able to make reliable predictions for different experiments and not only for the ones that were used to estimate the parameters. The goodness of the predictions is visualized plotting the output of the *in silico* experiment obtained with the model on top of the corresponding experimental data: G: Calnexin WT PAT6 Overexpression ^{35}S decay; H: Calnexin CA mutant ^{35}S decay; I: Calnexin CA mutant ^3H decay; L: Calnexin WT PAT6 Silencing ^{35}S decay.

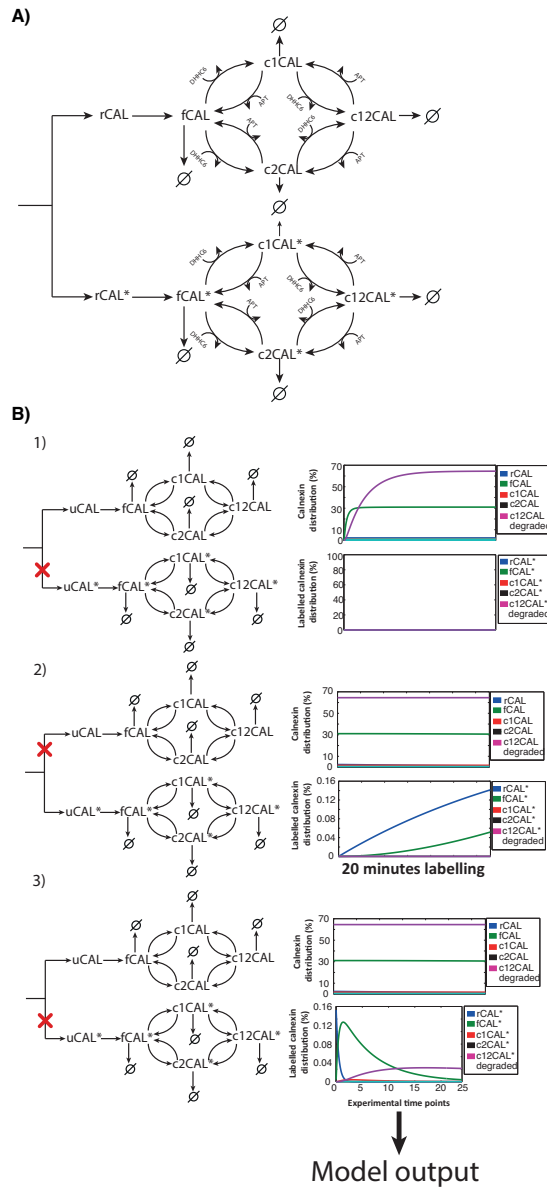


Figure S4.9. Schematic representation of the model and the procedure used for ^{35}S pulse-chase experiments simulation. (A) Model scheme used for ^{35}S labelling simulations. The labeled species are marked with *. (B) Qualitative representation of the method used to perform ^{35}S *in silico* labeling experiments. Each simulation is performed in 3 phases. 1) Steady state: the model of the endogenous calnexin is simulated until it reaches steady state. 2) Labeling: the synthesis term is switched towards the labeled protein (marked with *) to simulate the addition of the label for a time that matches the experiments. 3) Analysis: the synthesis term is switched again towards endogenous calnexin to simulate the removal of the labeling. The decay of the labeled species is observed at the same time point of the experiments.

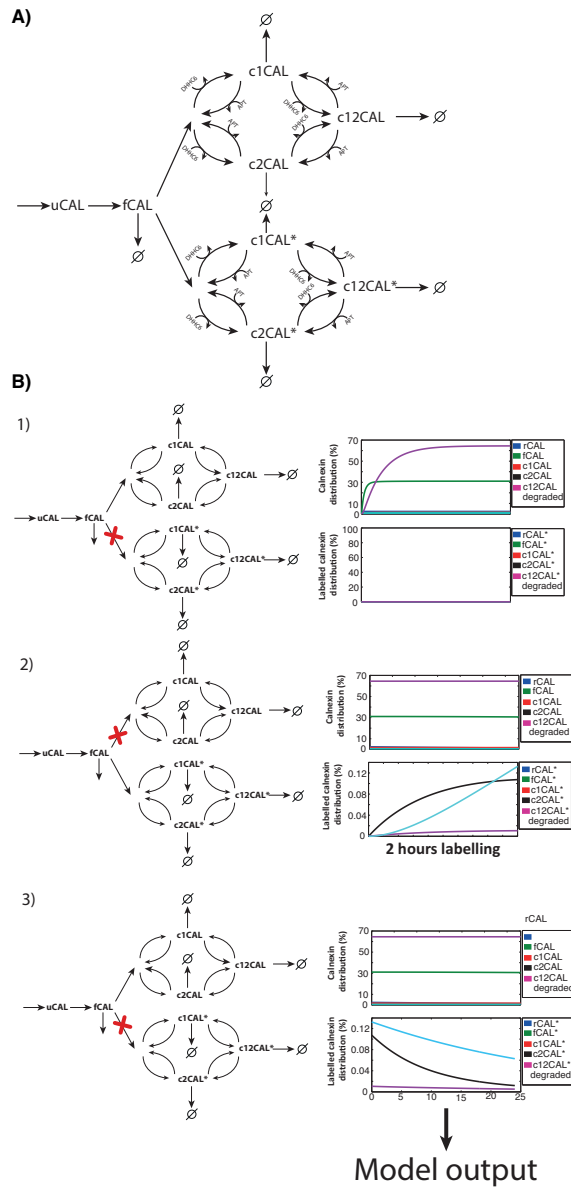


Figure S4.10. Schematic representation of the model and the procedure used for ^3H pulse-chase experiments simulation. (A) Model scheme used for ^3H labelling simulations. The labeled species are marked with *. (B) Qualitative representation of the steps used to perform ^3H in silico labelling experiments. Each simulation is performed in 3 phases. 1) Steady state: the model of the endogenous calnexin is simulated until it reaches steady state. 2) Labeling: the reactions of palmitate attachment are switched towards the labeled protein (marked with *) to simulate the addition of the ^3H palmitate for a time that matches the experiments. 3) Analysis: the reactions of palmitate attachment switched again towards endogenous calnexin to simulate the removal of the radiolabelled palmitate. The decay of the labeled species is observed at the same time point of the experiments.

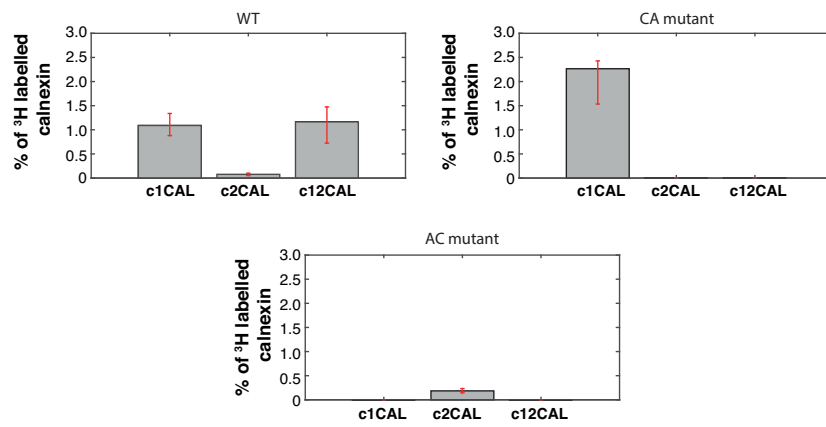


Figure S4.II. Distribution of the labelled species in the different palmitoylation states after 2 hours labeling with ³H palmitate. The relative amount of labelled protein with respect to the total pool of calnexin inside the cell (2 hours palmitate labelling) is also shown.

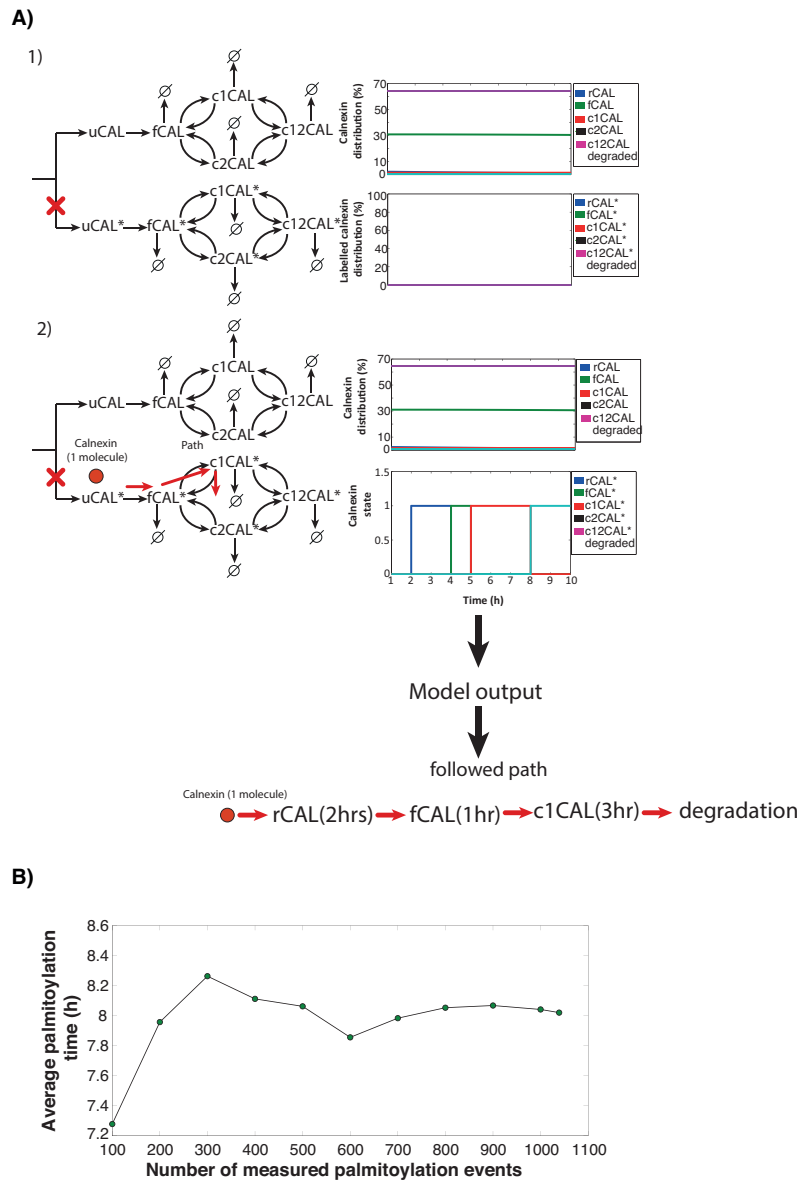


Figure S4.12. Qualitative representation of the procedure used for protein tracking with stochastic simulations, along with proof of convergence. (A) Each simulation is performed in 2 phases. 1) Steady state: the model is simulated until the pool of endogenous calnexin is in steady state. 2) Analysis: a single calnexin molecule is added to the pool of labeled protein (marked with *), and its route is recorded at each time step. Analysis of the route allows to calculate the time for going from synthesis to the double palmitoylation state. (B) The average palmitoylation time measured have converged to 8h after 1100 event of double palmitoylation tracked. We performed 5000 single molecule tracking simulations. During these simulations only 1100 molecule were able to get double palmitoylated before getting degraded. Here we reported the cumulative average palmitoylation time to show that the measured value converged to 8h.

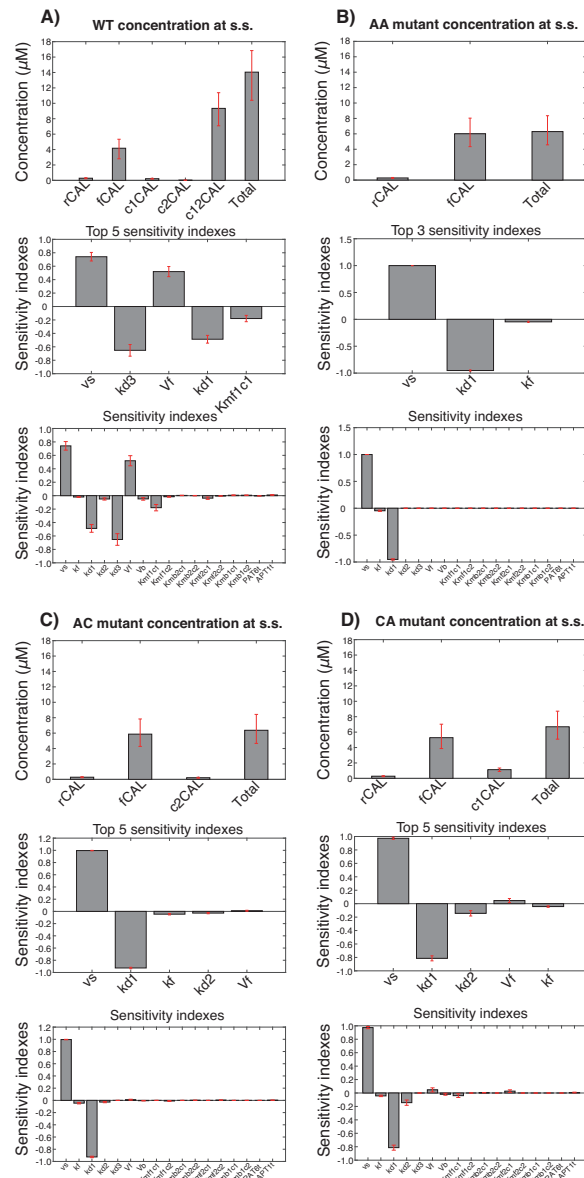


Figure S4.13. Steady state sensitivity analysis of calnexin WT and mutants. In order to determine to which parameters the steady state of calnexin is most sensitive we performed sensitivity analysis on WT and mutants.

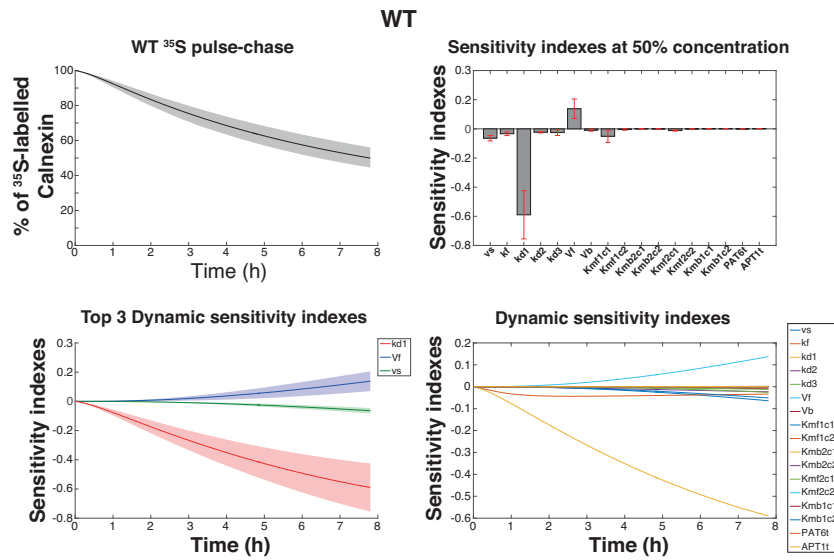


Figure S4.14. Half life sensitivity analysis of calnexin WT. Experimental data and the analysis of the model parameters have shown how palmitoylation affects stability of calnexin. Sensitivity analysis on *in-silico* ³⁵S experiments confirm that the protein half life is strongly influenced by the ability of calnexin of being palmitoylated.

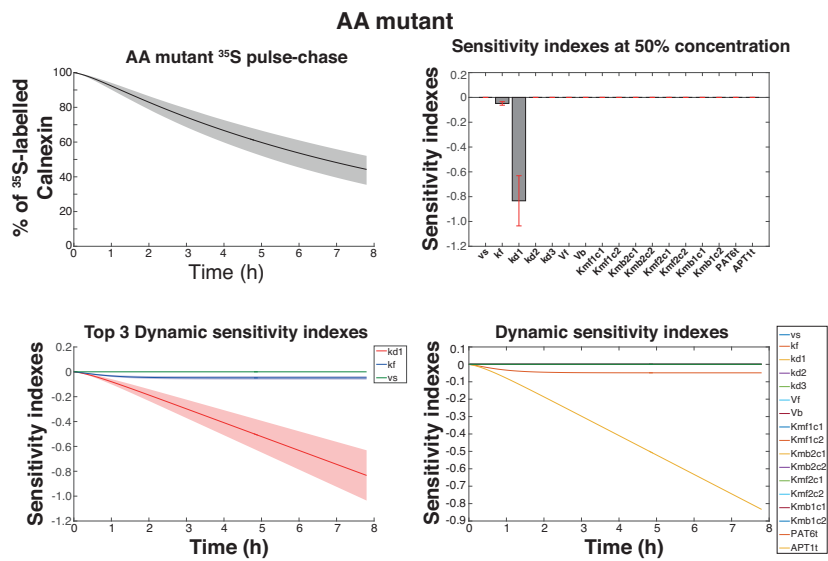


Figure S4.15. Half life sensitivity analysis of calnexin AA mutant. Sensitivity analysis on *in-silico* ³⁵S experiments on the AA calnexin mutant shows that if the protein can't be palmitoylated, it loses the capabilities to regulate its half-life. The stability of the protein become dependent from the degradation rate of the non palmitoylated state only.

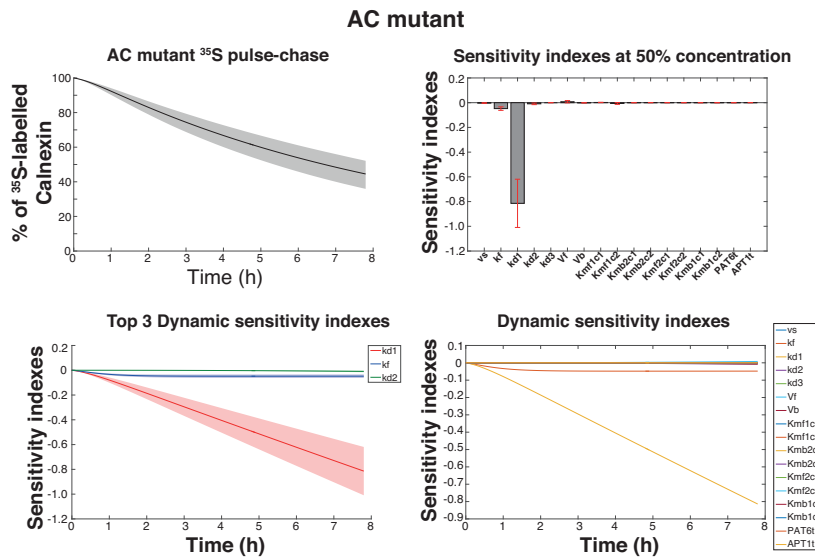


Figure S4.16. Half life sensitivity analysis of calnexin AC mutant. Sensitivity analysis on *in-silico* ^{35}S experiments on the AC calnexin mutant. Results show that a single palmitoylation site is not enough to increase the stability of the protein which remain dependent only from the degradation rate of the non palmitoylated state only.

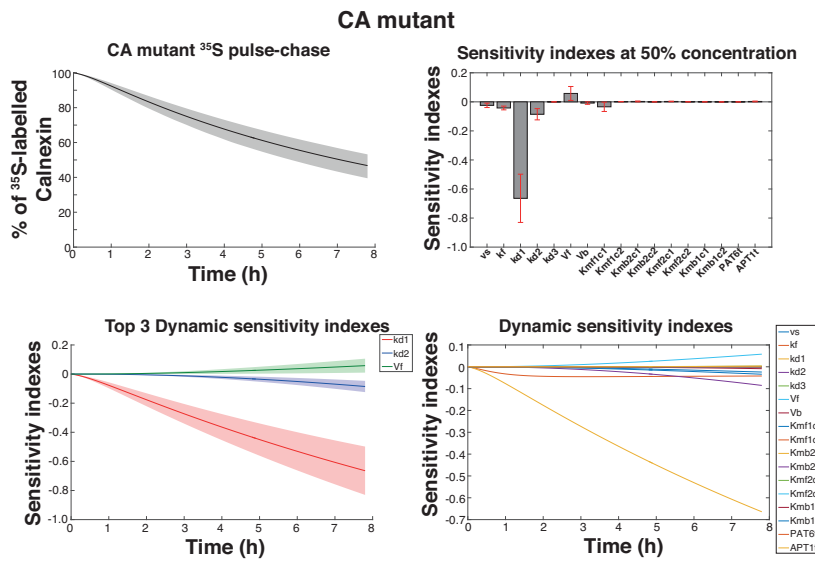


Figure S4.17. Half life sensitivity analysis of calnexin CA mutant. Sensitivity analysis on *in-silico* ^{35}S experiments on the CA calnexin mutant confirm that the presence of only 1 site isn't enough to be able to regulate the stability of the protein. If compared with S9 Fig, it is possible to see how the two sites are not equivalent, palmitoylation of the first site influence more the stability of calnexin.

Chapter 4: Model-Driven Understanding of Palmitoylation Dynamics: Regulated Acylation of The Endoplasmic Reticulum Chaperone Calnexin

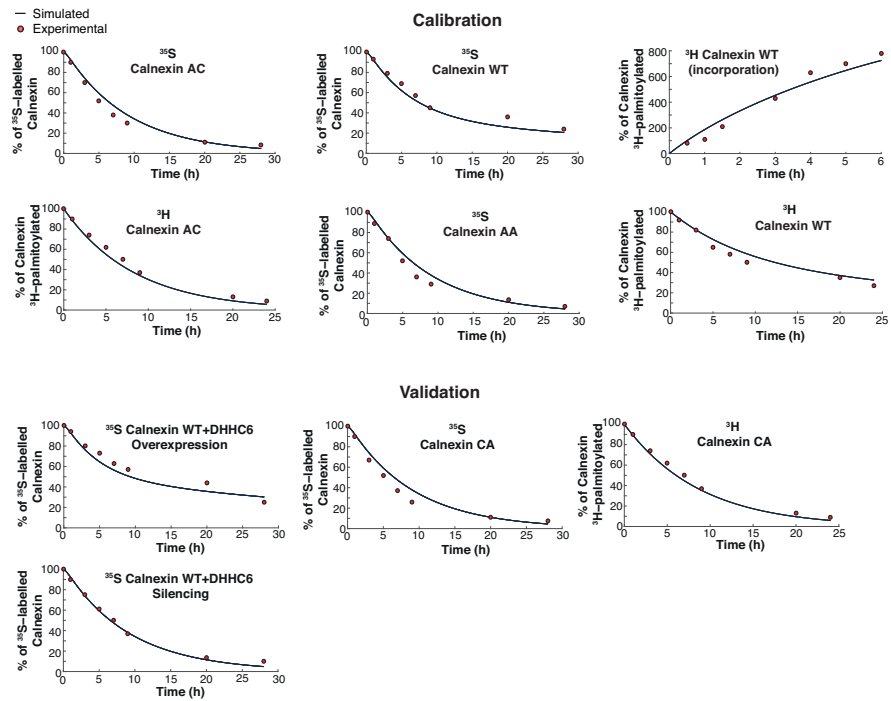


Figure S4.18. Fitting of the data used for calibration and validation after averaging the 383 sets of parameters estimated by the GA. In order to estimate a unique set of parameters for the stochastic simulations the 383 sets of parameter were averaged to obtain a single parameter set that was then converted to stochastic. The figure shows that the averaged set of parameters is still able to reproduce both the calibration and validation data.

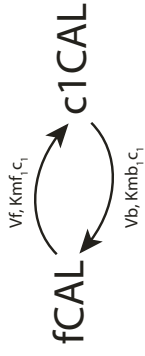
Table S4.2. Model reactions. The model of calnexin palmitoylation contains 14 different reactions, describing synthesis, folding, degradation, and the enzymatic reactions of calnexin palmitoylation/depalmitoylation. In the following table we describe in detail how the rates for those reactions are calculated.

Reaction	Forward rate	Reverse rate
1. Synthesis and folding		
$\rightarrow rCAL$	$v1 = vS$	-
$rCAL \rightarrow fCAL$	$v2 = kf \cdot rCAL$	-

$\frac{V_{f_2} \cdot C^{000}}{Km_{f_2} f_2}$ <p>2. Palmitoylation</p>		
$C^{011} \rightarrow \emptyset$ $fCAL \rightleftharpoons c_1 CAL$	$v3 = Vf \cdot \frac{\frac{fCAL}{Kmf_1 c_1}}{1 + \frac{fCAL}{Kmf_1 c_1} + \frac{fCAL}{Kmf_1 c_2} + \frac{c_1 CAL}{Kmf_2 c_1} + \frac{c_2 CAL}{Kmf_2 c_2} + \frac{PAT_6}{Kmf_1 c_1}}$	$v7 = Vb \cdot \frac{\frac{c_1 CAL}{Kmb_2 c_1}}{1 + \frac{c_1 CAL}{Kmb_2 c_1} + \frac{c_2 CAL}{Kmb_2 c_2} + \frac{c_{12} CAL}{Kmb_1 c_1} + \frac{c_{12} CAL}{Kmb_1 c_2} + \frac{APT_{1f}}{Kmb_2 c_1}}$
$fCAL \rightleftharpoons c_2 CAL$	$v4 = Vf \cdot \frac{\frac{fCAL}{Kmf_1 c_2}}{1 + \frac{fCAL}{Kmf_1 c_1} + \frac{fCAL}{Kmf_1 c_2} + \frac{c_1 CAL}{Kmf_2 c_1} + \frac{c_2 CAL}{Kmf_2 c_2} + \frac{PAT_6}{Kmf_1 c_2}}$	$v8 = Vb \cdot \frac{\frac{c_2 CAL}{Kmb_2 c_2}}{1 + \frac{c_1 CAL}{Kmb_2 c_1} + \frac{c_2 CAL}{Kmb_2 c_2} + \frac{c_{12} CAL}{Kmb_1 c_1} + \frac{c_{12} CAL}{Kmb_1 c_2} + \frac{APT_{1f}}{Kmb_2 c_2}}$
$c_1 CAL \rightleftharpoons c_{12} CAL$	$v5 = Vf \cdot \frac{\frac{c_1 CAL}{Kmf_2 c_1}}{1 + \frac{fCAL}{Kmf_1 c_1} + \frac{fCAL}{Kmf_1 c_2} + \frac{c_1 CAL}{Kmf_2 c_1} + \frac{c_2 CAL}{Kmf_2 c_2} + \frac{PAT_6}{Kmf_2 c_1}}$	$v9 = Vb \cdot \frac{\frac{c_{12} CAL}{Kmb_1 c_1}}{1 + \frac{c_1 CAL}{Kmb_2 c_1} + \frac{c_2 CAL}{Kmb_2 c_2} + \frac{c_{12} CAL}{Kmb_1 c_1} + \frac{c_{12} CAL}{Kmb_1 c_2} + \frac{APT_{1f}}{Kmb_1 c_1}}$

$c_2CAL \rightleftharpoons c_{12}CAL$	$v_6 = Vf \cdot \frac{\frac{c_2CAL}{Kmf_2c_2}}{1 + \frac{fCAL}{Kmf_1c_1} + \frac{fCAL}{Kmf_1c_2} + \frac{c_1CAL}{Kmf_2c_1} + \frac{c_2CAL}{Kmf_2c_2} + \frac{PAT_6}{Kmf_2c_2}}$	$v_{10} = Vb \cdot \frac{\frac{c_{12}CAL}{Kmb_1c_2}}{1 + \frac{c_1CAL}{Kmb_2c_1} + \frac{c_2CAL}{Kmb_2c_2} + \frac{c_{12}CAL}{Kmb_1c_1} + \frac{c_{12}CAL}{Kmb_1c_2} + \frac{APT_{1f}}{Kmb_1c_2}}$
3. Degradation		
$fCAL \rightarrow$	$v_{11} = kd_1 \cdot fCAL$	-
$c_1CAL \rightarrow$	$v_{12} = kd_2 \cdot c_1CAL$	-
$c_2CAL \rightarrow$	$v_{13} = kd_2 \cdot c_2CAL$	-
$c_{12}CAL \rightarrow$	$v_{14} = kd_3 \cdot c_{12}CAL$	-

Where for each palmitoylation and depalmitoylation step we define the kinetic parameters of the reaction as:



$$V_f = k_{catP6} \cdot PAT6_{tot}$$

$$K_{mf, c_1} = \frac{k_{c1P6ub} + k_{catP6}}{k_{c1P6b}}$$

$$V_b = k_{catAPT} \cdot APT1_{tot}$$

$$K_{mb, c_1} = \frac{k_{c1APTub} + k_{catAPT}}{k_{c1APTb}}$$

where k_{catP6} is the catalytic constant for DHHC6, while k_{catAPT} is the catalytic constant for APT1. The catalytic constants are assumed to be the same in all the modification steps. K_{c1P6b} and k_{c1P6ub} are the binding and unbinding rate constants of fCAL to DHHC6 respectively. K_{c1APTb} and $k_{c1APTub}$ are the binding and unbinding rate constants of c1CAL to APT1 respectively. In our model we assume that the binding and unbinding rate to DHHC6 and APT can be different for each state of the model.

Table S4.3. Model parameters. The output of GA is a set of optimal solutions, where a solution is a complete set of parameter needed to perform model simulations. From this set we extracted a sub-set of 382 solutions which obtained a GA score better than a set threshold for each objective. During the analysis the model was simulated for each set of parameters of the sub-set. We then reported in this paper the mean of the outputs along with the 1st and 3rd quartile of their distribution.

Parameter	Mean Value	±s.d.	Range
vs vs[C]/[T]	0.63	±0.17	[0.30-0.87]
kf_1 /[T]	2.31	±0.34	[1.14-2.97]
kd_1 /[T]	0.11	±0.03	[0.08-0.16]
kd_2 /[T]	0.07	±0.01	[0.04-0.09]
kd_3 /[T]	0.02	±5 · 10 ⁻³	[0.01-0.03]
Vf [C]/[T]	0.76	±0.24	[0.40-1.14]
Vb [C]/[T]	0.10	±0.03	[0.05-0.14]
Kmf_1c_1 [C]	7.06	±0.83	[5.74-10.87]
Kmf_1c_2 [C]	87.07	±17.61	[51.88-121.46]
Kmb_2c_1 [C]	0.55	±0.16	[0.09-0.96]
Kmb_2c_2 [C]	0.51	±0.20	[0.03-0.89]
Kmf_2c_1 [C]	0.43	±0.08	[0.23-0.54]
Kmf_2c_2 [C]	0.31	±0.06	[0.19-0.47]
Kmb_1c_1 [C]	43.68	±6.80	[4.98-86.77]
Kmb_1c_2 [C]	38.60	±21.82	[1.94-88.13]
PAT_6 [C]	0.01	±1 · 10 ⁻³	[7 · 10 ⁻³ -0.01]
APT_1 [C]	0.87	±0.16	[0.58-1.32]

Table S4.4. Mass balance equations. In this model calnexin can exist in 5 different states: unfolded (rCAL), folded (fCAL), palmitoylated only on the first site (c1CAL) or the second (c2CAL), or dually palmitoylated (c12CAL). The following table describes the mass balance for each of these species. The rates of the mass balance of each state are described in detail in S1 Table.

Mass balance
$\frac{drCAL}{dt} = v_1 - v_2$
$\frac{dfCAL}{dt} = v_2 + v_7 + v_8 - v_{11} - v_3 - v_4$
$\frac{dc_1CAL}{dt} = v_3 + v_{10} - v_7 - v_{12} - v_5$
$\frac{dc_2CAL}{dt} = v_4 + v_{10} - v_8 - v_{13} - v_6$
$\frac{dc_{12}CAL}{dt} = v_5 + v_6 - 2 \cdot v_{10} - v_{14}$

Chapter 4: Model-Driven Understanding of Palmitoylation Dynamics: Regulated Acylation of
 The Endoplasmic Reticulum Chaperone Calnexin

Table S4.5. Stoichiometric matrix used for stochastic simulations. In the following table we define the stoichiometry and the directionality of the reactions of the model. Each reaction has a directionality that define which are the reagents and which are the products. In this table each line represents a model reaction, while in the columns we find all the states of the model. For each reaction the states of the model that take part to the reaction as reagents are marked with -1, while the states that participate as products are marked with 1. The matrix that is formed in this way allow to attribute the correct directionality to model reactions during the calculation of the mass balance for each state of the model.

	rCAL	fcAL	c1CAL	c2CAL	c12CAL	rCALI	fcALI	c1CALI	c2CALI	c12CALI	fcALIddeg	c1CALIddeg	c2CALIddeg	c12CALIddeg
0->rCAL	1	0	0	0	0	0	0	0	0	0	0	0	0	0
rCAL->fcAL	-1	1	0	0	0	0	0	0	0	0	0	0	0	0
fcAL->0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0
fcAL->c1CAL	0	-1	1	0	0	0	0	0	0	0	0	0	0	0
fcAL->c2CAL	0	-1	0	1	0	0	0	0	0	0	0	0	0	0
c1CAL->fcAL	0	1	-1	0	0	0	0	0	0	0	0	0	0	0
c1CAL->c12CAL	0	0	-1	0	1	0	0	0	0	0	0	0	0	0
c1CAL->0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0
c2CAL->fcAL	0	1	0	-1	0	0	0	0	0	0	0	0	0	0
c2CAL->c12CAL	0	0	0	-1	1	0	0	0	0	0	0	0	0	0
c2CAL->0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0
c12CAL->c1CAL	0	0	1	0	-1	0	0	0	0	0	0	0	0	0
c12CAL->c2CAL	0	0	0	1	-1	0	0	0	0	0	0	0	0	0
c12CAL->0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0
0->rCALI	0	0	0	0	0	1	0	0	0	0	0	0	0	0
rCALI->fcALI	0	0	0	0	0	-1	1	0	0	0	0	0	0	0
fcALI->0	0	0	0	0	0	0	-1	0	0	0	1	0	0	0
fcALI->c1CALI	0	0	0	0	0	0	-1	1	0	0	0	0	0	0
fcALI->c2CALI	0	0	0	0	0	0	0	-1	1	0	0	0	0	0
c1CALI->fcALI	0	0	0	0	0	0	1	-1	0	0	0	0	0	0
c1CALI->c12CALI	0	0	0	0	0	0	0	-1	1	0	0	0	0	0
c1CALI->0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0
c2CALI->fcALI	0	0	0	0	0	0	1	0	-1	0	0	0	1	0
c2CALI->c12CALI	0	0	0	0	0	0	0	0	-1	1	0	0	0	0
c2CALI->0	0	0	0	0	0	0	0	0	-1	0	0	0	1	0
c12CALI->c1CALI	0	0	0	0	0	0	1	0	0	-1	0	0	0	0
c12CALI->c2CALI	0	0	0	0	0	0	0	0	1	-1	0	0	0	0
c12CALI->0	0	0	0	0	0	0	0	0	0	0	-1	0	0	1

Table 4.6. Propensity function used for stochastic simulations. In the first column of the table, each line describes a reaction of the model. To each reaction is associated a rate, in the second column, that describes the probability of that reaction to happen at each time step of the stochastic simulation.

Reaction	Propensity
$\rightarrow rCAL$	νs
$rCAL \rightarrow fCAL$	$k_f \cdot rCAL$
$fCAL \rightarrow$	$k_{d_1} \cdot fCAL$
$fCAL \rightarrow c_1CAL$	$Vf \cdot \frac{fCAL}{Kmf_1c_1} + \frac{fCAL}{Kmf_1c_1} + \frac{c_2CAL}{Kmf_2c_2} + \frac{fCAL^*}{Kmf_1c_1} + \frac{fCAL^*}{Kmf_1c_1} + \frac{c_1CAL^*}{Kmf_2c_1} + \frac{fCAL^*}{Kmf_2c_2} + \frac{PAT6t}{Kmf_1c_1} + \frac{c_2CAL^*}{Kmf_2c_2} + \frac{PAT6t}{Kmf_1c_1}$
$fCAL \rightarrow c_2CAL$	$Vf \cdot \frac{fCAL}{Kmf_1c_2} + \frac{fCAL}{Kmf_1c_2} + \frac{c_1CAL}{Kmf_2c_1} + \frac{c_2CAL}{Kmf_2c_2} + \frac{fCAL^*}{Kmf_1c_2} + \frac{fCAL^*}{Kmf_1c_2} + \frac{c_1CAL^*}{Kmf_2c_1} + \frac{fCAL^*}{Kmf_2c_2} + \frac{PAT6t}{Kmf_1c_2} + \frac{c_2CAL^*}{Kmf_2c_2} + \frac{PAT6t}{Kmf_1c_2}$
$c_1CAL \rightarrow fCAL$	$Vb \cdot \frac{c_1CAL}{Kmb_2c_1} + \frac{c_2CAL}{Kmb_2c_2} + \frac{c_1CAL}{Kmb_1c_1} + \frac{c_2CAL}{Kmb_1c_2} + \frac{c_1CAL^*}{Kmb_2c_1} + \frac{c_2CAL^*}{Kmb_2c_2} + \frac{c_1CAL^*}{Kmb_1c_1} + \frac{c_2CAL^*}{Kmb_1c_2} + \frac{APT1t}{Kmb_2c_1} + \frac{APT1t}{Kmb_2c_2}$

$c_1CAL \rightarrow c_{12}CAL$	$Vf \cdot \frac{c_1CAL}{Kmf_2c_1} \cdot \frac{fCAL}{1 + \frac{fCAL}{Kmf_1c_1} + \frac{fCAL}{Kmf_1c_2} + \frac{c_1CAL}{Kmf_2c_1} + \frac{c_2CAL}{Kmf_2c_2} + \frac{fCAL^*}{Kmf_1c_1} + \frac{fCAL^*}{Kmf_1c_2} + \frac{c_1CAL^*}{Kmf_2c_1} + \frac{c_2CAL^*}{Kmf_2c_2} + \frac{PAT6t}{Kmf_2c_1}}$
$c_1CAL \rightarrow$	$kd_2 \cdot c_1CAL$
$c_2CAL \rightarrow fCAL$	$Vb \cdot \frac{c_2CAL}{Kmb_2c_2} \cdot \frac{c_1CAL}{1 + \frac{c_1CAL}{Kmb_2c_1} + \frac{c_2CAL}{Kmb_2c_2} + \frac{c_1CAL}{Kmb_1c_1} + \frac{c_2CAL}{Kmb_1c_2} + \frac{c_1CAL^*}{Kmb_2c_1} + \frac{c_2CAL^*}{Kmb_2c_2} + \frac{c_1CAL^*}{Kmb_1c_1} + \frac{c_2CAL^*}{Kmb_1c_2} + \frac{APT1t}{Kmb_2c_2}}$
$c_2CAL \rightarrow c_{12}CAL$	$Vf \cdot \frac{c_2CAL}{Kmf_2c_2} \cdot \frac{fCAL}{1 + \frac{fCAL}{Kmf_1c_1} + \frac{fCAL}{Kmf_1c_2} + \frac{c_2CAL}{Kmf_2c_1} + \frac{c_2CAL}{Kmf_2c_2} + \frac{fCAL^*}{Kmf_1c_1} + \frac{fCAL^*}{Kmf_1c_2} + \frac{c_2CAL^*}{Kmf_2c_1} + \frac{c_2CAL^*}{Kmf_2c_2} + \frac{PAT6t}{Kmf_2c_2}}$
$c_2CAL \rightarrow$	$kd_2 \cdot c_2CAL$
$c_{12}CAL \rightarrow c_1CAL$	$Vb \cdot \frac{c_{12}CAL}{Kmb_1c_1} \cdot \frac{c_1CAL}{1 + \frac{c_1CAL}{Kmb_2c_1} + \frac{c_2CAL}{Kmb_2c_2} + \frac{c_1CAL}{Kmb_1c_1} + \frac{c_2CAL}{Kmb_1c_2} + \frac{c_1CAL^*}{Kmb_2c_1} + \frac{c_2CAL^*}{Kmb_2c_2} + \frac{c_1CAL^*}{Kmb_1c_1} + \frac{c_2CAL^*}{Kmb_1c_2} + \frac{APT1t}{Kmb_1c_1}}$

$c_1CAL^* \rightarrow fCAL^*$	$Vb \cdot \frac{\frac{c_1CAL^*}{Kmb_2c_1}}{1 + \frac{c_1CAL}{Kmb_2c_1} + \frac{c_2CAL}{Kmb_1c_1} + \frac{c_1CAL}{Kmb_2c_1} + \frac{c_2CAL}{Kmb_1c_1} + \frac{c_1CAL^*}{Kmb_2c_1} + \frac{c_2CAL^*}{Kmb_1c_1} + \frac{c_1CAL^*}{Kmb_2c_1} + \frac{c_2CAL^*}{Kmb_1c_1} + \frac{APT1r}{Kmb_2c_1}}$
$c_1CAL^* \rightarrow c_{12}CAL^*$	$Vf \cdot \frac{\frac{c_1CAL^*}{Kmf_2c_1}}{1 + \frac{fCAL}{Kmf_1c_1} + \frac{fCAL}{Kmf_2c_1} + \frac{c_1CAL}{Kmf_2c_1} + \frac{c_2CAL}{Kmf_1c_1} + \frac{fCAL^*}{Kmf_1c_1} + \frac{fCAL^*}{Kmf_2c_1} + \frac{c_1CAL^*}{Kmf_1c_1} + \frac{c_2CAL^*}{Kmf_2c_1} + \frac{PAT6r}{Kmf_2c_1}}$
$c_1CAL^* \rightarrow$	$kd_2 \cdot c_1CAL^*$
$c_2CAL^* \rightarrow fCAL^*$	$Vb \cdot \frac{\frac{c_2CAL^*}{Kmb_2c_2}}{1 + \frac{c_1CAL}{Kmb_2c_1} + \frac{c_2CAL}{Kmb_2c_2} + \frac{c_1CAL}{Kmb_1c_1} + \frac{c_2CAL}{Kmb_1c_2} + \frac{c_1CAL^*}{Kmb_2c_1} + \frac{c_2CAL^*}{Kmb_2c_2} + \frac{c_1CAL^*}{Kmb_1c_1} + \frac{c_2CAL^*}{Kmb_1c_2} + \frac{APT1r}{Kmb_2c_2}}$
$c_2CAL^* \rightarrow c_{12}CAL^*$	$Vf \cdot \frac{\frac{c_2CAL^*}{Kmf_2c_2}}{1 + \frac{fCAL}{Kmf_1c_1} + \frac{fCAL}{Kmf_2c_2} + \frac{c_1CAL}{Kmf_2c_1} + \frac{c_2CAL}{Kmf_1c_2} + \frac{fCAL^*}{Kmf_1c_1} + \frac{fCAL^*}{Kmf_2c_2} + \frac{c_1CAL^*}{Kmf_1c_1} + \frac{c_2CAL^*}{Kmf_2c_2} + \frac{PAT6r}{Kmf_2c_2}}$
$c_2CAL^* \rightarrow$	$kd_2 \cdot c_2CAL^*$

$c_{12}CAL^* \rightarrow c_1CAL^*$	$Vb \cdot \frac{\frac{c_{12}CAL^*}{Kmb_1c_1}}{1 + \frac{c_1CAL}{Kmb_2c_1} + \frac{c_2CAL}{Kmb_2c_2} + \frac{c_{12}CAL}{Kmb_1c_1} + \frac{c_1CAL^*}{Kmb_2c_1} + \frac{c_2CAL^*}{Kmb_2c_2} + \frac{c_{12}CAL^*}{Kmb_1c_1} + \frac{c_1CAL^*}{Kmb_1c_2} + \frac{c_2CAL^*}{Kmb_1c_2} + \frac{APT}{Kmb_1c_1}}$
$c_{12}CAL^* \rightarrow c_2CAL^*$	$Vb \cdot \frac{\frac{c_{12}CAL^*}{Kmb_1c_2}}{1 + \frac{c_1CAL}{Kmb_2c_1} + \frac{c_2CAL}{Kmb_2c_2} + \frac{c_{12}CAL}{Kmb_1c_2} + \frac{c_1CAL^*}{Kmb_2c_1} + \frac{c_2CAL^*}{Kmb_2c_2} + \frac{c_{12}CAL^*}{Kmb_1c_1} + \frac{c_1CAL^*}{Kmb_1c_2} + \frac{c_2CAL^*}{Kmb_1c_2} + \frac{APT}{Kmb_1c_2}}$
$c_{12}CAL^* \rightarrow$	$kd_3 \cdot c_{12}CAL^*$

Table S4.7. Parameters used for stochastic simulations. The following parameters were obtained through the conversion of the deterministic parameters estimated by the GA (see “Conversion of deterministic parameters to stochastic”).

Parameter	Value	Units
vs	18450	molecules/h
kf	2.31	1/h
kd_1	0.11	1/h
kd_2	0.07	1/h
kd_3	0.02	1/h
Vf	22314	molecules/h
Vb	2933	molecules/h
Kmf_1c_1	206110	molecules
Kmf_1c_2	2543152	molecules
Kmb_2c_1	16013	molecules
Kmb_2c_2	14944	molecules
Kmf_2c_1	12559	molecules
Kmf_2c_2	8987	molecules
Kmb_1c_1	1275827	molecules
Kmb_1c_2	1127490	molecules
PAT_6	380	molecules
APT_1	25372	molecules

Chapter 5 - DHHC16-mediated control of DHHC6 activity: dynamics of a palmitoylation cascade

5.1 Introduction

Cells constantly interact and respond to their environment. This requires tight control of protein function in time and in space, which largely occurs through the reversible post-translational modification, such as phosphorylation ubiquitination and S-palmitoylation. The latter consist in the addition on an acyl chain, generally C16, to cytosolic cysteine residues, thereby increasing the hydrophobicity of the protein. As a result palmitoylation allows proteins to interact with membranes or specific membrane domains, it may affect their conformation, trafficking, stability and function [47,115–117]. The acyl chain is attached to the protein via a thioester bond through the action of palmitoyltransferases of the DHHC family [47,115–117]. These are multispinning transmembrane proteins harboring their active site in the cytoplasm.

The list of proteins undergoing palmitoylation is ever increasing [16] and the modification is found to be important in numerous key cellular processes including the control of neuronal kinase activity [118], of specific forms of autophagy, of EGF signaling [119], of metastasis in cancer [120].

While novel roles and targets of palmitoylation are constantly reported, little is known about the regulation of this post-translational modification. Here we focused on one of the 23 human palmitoyltransferase, DHHC6, which localizes to the ER and controls

a panel of key ER substrates such as the ER chaperone calnexin [58], the ER E3 ligase gp78 [121], the IP3 receptor [122] as well as cell surface proteins such as the transferrin receptor [123]. For each of these proteins, palmitoylation controls stability, localization, trafficking and/or function.

As most DHHC enzymes, DHHC6 is a tetra-spanning membrane protein. It is characterized by a short N-terminal extension and a long C-terminal tail composed of an approx. 100 residue domain of unknown structure followed by a variant SH3_2 domain (Figure 5.1A, Figure S5.7).

At the very far end it contains a KKNR motif which when transferred to the DHHC5 enzyme leads to its relocation from the Golgi to the ER [124], suggesting it contributes to the ER localization of DHHC6. The function of the C-terminal tail, and in particular the SH3_2 domain is unknown. Here we show that this SH3_2 domain undergoes dynamic palmitoylation and depalmitoylation on three different sites, and that these events are essential for proper localization of the enzyme, for its stability and for its function. We identify both the palmitoyltransferase and the acyl protein thioesterase involved in this modification. Using mathematical modelling we probe the complexity of this system, wherein DHHC6, due to the presence of three sites, can exist as 8 different species with different stabilities and activities. Importantly, the model provided testable predictions regarding the importance of having three acylation sites. Predictions and experiments indicate that palmitoylation on multiple sites allows the tight control of both DHHC6 activity and abundance.

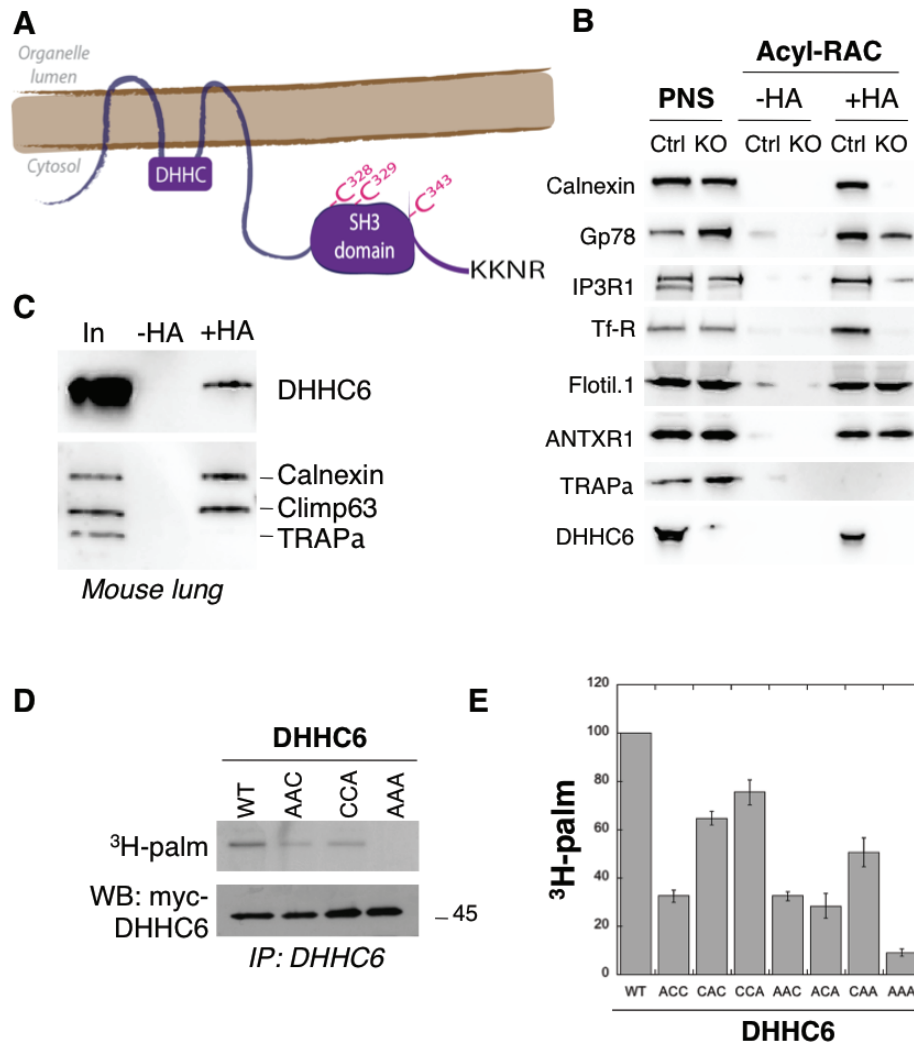


Figure 5.1. DHHC6 can undergo palmitoylation on three cysteines of its SH3 domain. A. Schematic representation of DHHC6 enzyme. Palmitoylated cysteines 328,329 and 343 are in red. DHHC and SH3 cytosolic domains are in blue. The potential ER retention motif is KKNR. B. Analysis of protein acylation in HAP WT control cells (Ctrl) versus HAP KO for DHHC6 (KO). HAP cell membranes were recovered by centrifugation and incubated with MMTS and then with hydroxylamine (+HA) or with TRIS (-HA) together with free thiol group binding beads. Eluted fractions were analyzed by immunoblotting with the indicated antibodies. The input fraction (PNS) was loaded as 1/10 this amount. C. DHHC6 acylation in mouse tissues. 400µg total proteins extracted from mouse lung were incubated with MMTS and then with hydroxylamine (+HA) or with TRIS (-HA) together with free thiol group binding beads. Eluted fractions were analyzed by immunoblotting with the indicated antibodies. The input fraction (In) was loaded as 1/10 this amount. D. Palmitoylation of DHHC6 cysteine mutants. HeLa cells were transfected with plasmids encoding WT or the indicated mutant Myc-tagged DHHC6 constructs for 24h. Cells were then metabolically labeled 2 hours at 37°C with ³H-palmitic acid. Proteins were extracted, immunoprecipitated with Myc antibodies and subjected to SDS-PAGE and analyzed by autoradiography (³H-palm), quantified using the Typhoon Imager or by immunoblotting with the indicated antibodies. E. Quantification of ³H-palmitic acid incorporation into DHHC6 constructs. Quantified values were normalized to protein expression level. The calculated value of ³H-palmitic acid incorporation into WT DHHC6 was set to 100% and all mutants were expressed relative to this (n=4, error bars represent standard deviation).

5.2 Results

5.2.1 Palmitoylation of the DHHC6 SH3 domain

DHHC6 KO cells were generated using the CRISPR-cas9 system in the near haploid cell line HAPI. Using the Acyl-RAC capture method to isolate palmitoylated proteins, we verified that the ER chaperone calnexin [58], the E3 ligase gp78 [121], the IP3 receptor [122] and the transferrin receptor [123] are indeed DHHC6 targets (Figure 5.1B). Interestingly calnexin and the transferrin receptor were no longer palmitoylated in the HAPI DHHC6 KO cells, confirming that they are exclusively modified by DHHC6, while palmitoylation of the IP3 receptor and gp78 was reduced but not abolished indicating that they can also be modified by other palmitoyltransferases (Figure 5.1B) consistent with previous findings [121,122]. As negative controls, we tested flotilin2, a known target of DHHC5 [125] and the anthrax toxin receptor 1 modified by a yet to be determined DHHC enzymes [46].

The acyl-RAC analysis indicated that DHHC6 itself might be palmitoylated (Figure 5.1B) in HAPI cells, as also found upon acyl-RAC analysis of different mouse tissues (shown for lung tissue in Figure 5.1C). Calnexin and the ER shaping protein climp63 [126] were used as positive controls and the cysteine-less protein Trap α as negative control. A large scale proteomics analysis also reported palmitoylation of DHHC6 on Cys-328, Cys-329 and Cys-343, which all localize to the SH3 domain [127] and are conserved in vertebrates (Figure S5.7). We generated single, double and triple cysteine-to-alanine mutants and monitored the incorporation of ^3H -palmitate during 2hrs. Palmitoylation of WT DHHC6 was readily detected (Figure 5.1D). The signal was reduced upon mutation of any of the cysteines, and especially Cys-328 (Figure 5.1E). The triple AAA mutant showed only background levels of ^3H -palmitate labeling. Thus the SH3 domain of DHHC6 can undergo palmitoylation on all three cysteines.

5.2.2 The DHHCl6-DHHC6 palmitoylation cascade

We next investigated whether DHHC6 was undergoing auto-palmitoylation, in cis or trans. We generated a construct to express an inactive enzyme by deleting the DHHC motif (Δ DHHC) as well as a stable cell line expressing an shRNA against DHHC6 (Figure S5.8). WT and Δ DHHC DHHC6 both underwent palmitoylation when transiently expressed in control or shRNA DHHC6 cells (Figure S5.8), showing that DHHC6 does not undergo autopalmitoylation but must be modified by another enzyme. To identify this enzyme, we performed ^3H -palmitate DHHC6 labeling experiments upon siRNA silencing of each of the 23 DHHC enzymes. Only silencing of DHHCl6 led to a major, 60%, drop in signal intensity (Figure 5.2A). ^3H -palmitate incorporation was also reduced for the AAC and CCA mutants (Figure S5.8A), showing that all three sites are modified by the same enzyme. A screen by over-expression of DHHC enzymes also pointed to DHHCl6 as the responsible enzyme (Figure 5.2B and Figure S5.8B).

The final confirmation was obtained using CRISPR-Cas9-generated DHHCl6 KO HAPI cells. WT and Δ DHHC DHHC6 incorporated ^3H -palmitate when expressed in the control and the DHHC6 KO cells, but not in DHHCl6 KO cells (Figure 5.2C). Consistent with these findings, co-immunoprecipitation experiments following transient over-expression of myc-tagged DHHC6 and FLAG-tagged DHHCl6 confirm that the two enzymes can interact (Figure S5.9A).

Altogether these experiments show that DHHC6 can be palmitoylated on all three of its SH3 cysteine residues by DHHCl6. DHHCl6 itself is not palmitoylated as shown both by ^3H -palmitate incorporation and Acyl-RAC (Figure S5.9BC).

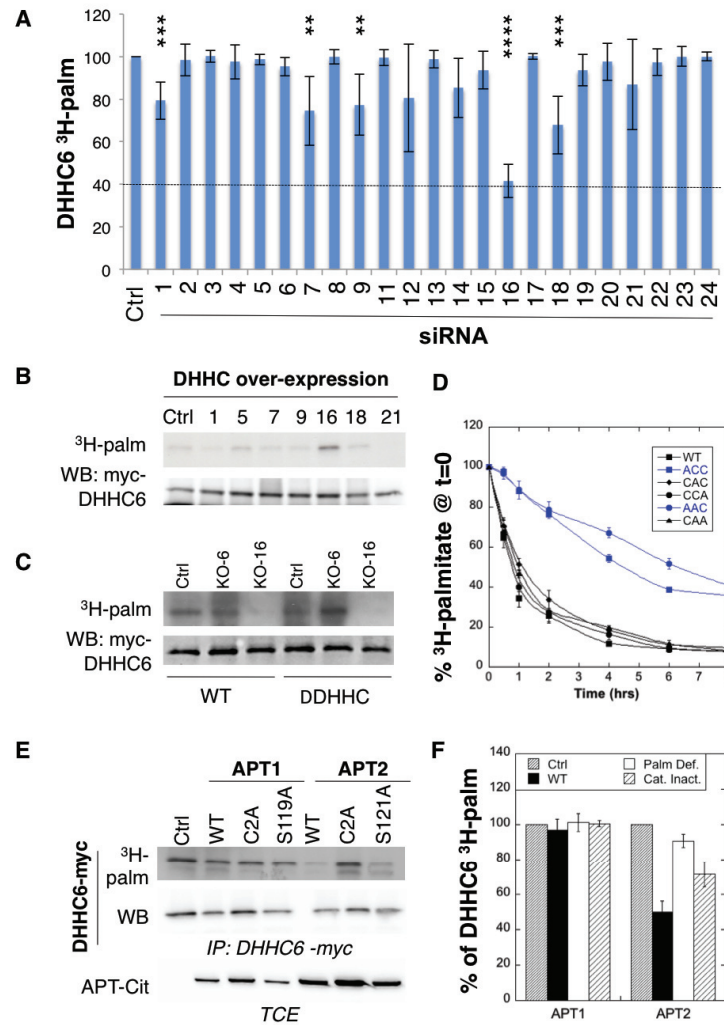


Figure 5.2. DHHC6 is palmitoylated by DHHC16. **A.** Identification of the DHHC6 palmitoyltransferase by siRNA screening of DHHC enzymes. HeLa cells were transfected with siRNA silencing indicated DHHC for 72h and with WT myc-tagged DHHC6 construct for the last 24h. Cells were then metabolically labeled 2 hours at 37°C with ³H-palmitic acid. Proteins were extracted, immunoprecipitated with myc antibodies and subjected to SDS-PAGE and analyzed by autoradiography, quantified using the Typhoon Imager or by immunoblotting with myc antibodies. ³H-palmitic acid incorporation into DHHC6 constructs were quantified and normalized to protein expression level. The calculated value of ³H-palmitic acid incorporation into DHHC6 was set to 100% for a non relevant siRNA (Ctrl) and all siRNA were expressed relative to this (n=6). **B.** Identification of the DHHC6 palmitoyltransferase by DHHC over-expression. HeLa cells were transfected with indicated DHHC constructs and with WT myc-tagged DHHC6 construct for 24h. Cells were then metabolically labeled 2 hours at 37°C with ³H-palmitic acid. Proteins were extracted, immunoprecipitated with myc antibodies and subjected to SDS-PAGE and analyzed by autoradiography (³H-palm) or by immunoblotting with myc antibodies. **C.** Analysis of DHHC6 acylation in HAP WT control cells (Ctrl) versus HAP KO for DHHC6 (KO) or KO for DHHC16. HAP cells were transfected with WT myc-tagged DHHC6 construct for 24h. Cells were then metabolically labeled 2 hours at 37°C with ³H-palmitic acid. Proteins were extracted, immunoprecipitated with myc antibodies and subjected to SDS-PAGE and analyzed by autoradiography (³H-palm) or by immunoblotting with myc antibodies. **D.** Palmitoylation decay of WT or mutant DHHC6. HeLa cells were transfected with plasmids encoding WT or the indicated mutant MYC-tagged DHHC6 constructs for 24h. Cells were then metabolically labeled 2 hours at 37°C with ³H-palmitic acid, washed and incubated with complete medium for different hours. Proteins were extracted, immunoprecipitated with myc antibodies and subjected to SDS-PAGE and analyzed by autoradiography, quantified using the Typhoon Imager or by immunoblotting with anti-myc antibodies. ³H-palmitic acid incorporation into different DHHC6 constructs were quantified for each times, normalized to protein expression level. The calculated value of ³H-palmitic acid incorporation into DHHC6 was set to 100% for t=0 after the 2 hours pulse and all different times of chase with complete medium were expressed relative to this. N=3. **E.** DHHC6 palmitoylation in presence of APT overexpressed. HeLa cells were transfected with plasmids encoding WT myc-DHHC6 and the indicated mutant citrin-tagged of APT1 or APT2 constructs for 24h. Cells were then metabolically labeled 2 hours at 37°C with ³H-palmitic acid.

Proteins were extracted, immunoprecipitated with myc antibodies and subjected to SDS-PAGE and analyzed by autoradiography (³H-palm), quantified using the Typhoon Imager or by immunoblotting with anti-myc antibodies. F. Quantification of ³H-palmitic acid incorporation into DHHC6 constructs. Quantified values were normalized to protein expression level. The calculated values of ³H-palmitic acid incorporation into WT DHHC6 were set to 100% when cells were cotransfected with control plasmid (Ctrl) and all APT mutants were expressed relative to this. N=6.

5.2.3 Rapid APT2-mediated DHHC6 depalmitoylation

Palmitoylation is a reversible modification and thus has the potential to be dynamic. To analyze palmitate turnover on DHHC6, we performed ³H-palmitate pulse-chase experiments. Following a 2hr pulse, 50% of the palmitate was released from DHHC6 in less than one hour, indicating rapid turnover of the acyl chains (Figure 5.2D). We performed similar experiments on the various single and double cysteine mutants and found that rapid turnover required the presence of Cys-328. In its absence, palmitate release rates were drastically reduced, the half-life of DHHC6-bound palmitate increasing to more than 4 hrs (Figure 5.2D).

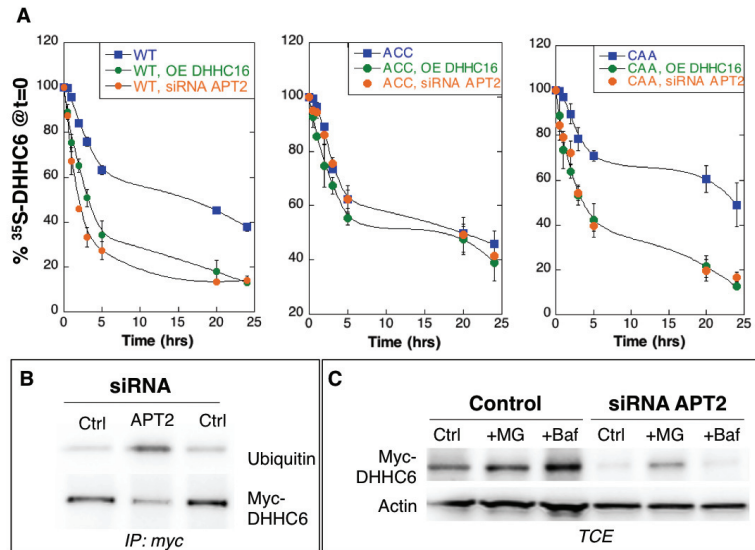


Figure 5.3. A. Degradation kinetics of DHHC6. HeLa cells were transfected with plasmids encoding WT flag-DHHC16, or WT myc-DHHC6 or cysteine mutants myc-DHHC6 constructs for 24h after 48h transfection with siRNA APT2 or with control siRNA. HeLa cells were incubated 20min pulse with ³⁵S-methionin/cysteine at 37°C, washed and further incubated for different times at 37°C in complete medium. DHHC6 were immunoprecipitated and subjected to SDS-PAGE and analyzed by autoradiography, quantified using the Typhoon Imager, and western blotting with anti-myc antibodies. ³⁵S-methionin/cysteine incorporation into different DHHC6 constructs were quantified for each times, normalized to protein expression level. The calculated value of ³⁵S-methionin/cysteine incorporation into DHHC6 was set to 100% for t=0 after the 20min pulse and all different times of chase with complete medium were expressed relative to this. N=3. B. Ubiquitin dependent degradation of DHHC6. HeLa cells were transfected with plasmids encoding WT myc-DHHC6 constructs for 24h after 48h transfection with control (Ctrl) or APT2

siRNA. Proteins were extracted; DHHC6 was immunoprecipitated, subjected to SDS-PAGE and then analyzed by immunoblotting with anti-ubiquitin or anti-myc antibodies. C. Proteosomal degradation of DHHC6. HeLa cells were transfected with plasmids encoding WT myc-DHHC6 constructs for 24h after 48h transfection with control or APT2 siRNA. Cells were treated 4 hours with 10uM MG132 or with 100 nM Bafilomycin A before proteins were extracted; 40ug of total extract were subjected to SDS-PAGE and then analyzed by immunoblotting with anti-actin as equal loading or with anti-myc antibodies.

Depalmitoylation is an enzymatic reaction that is mediated by poorly characterized Acyl Protein Thioesterases (APT). We tested the involvement of APT1 and APT2 [47], which are themselves palmitoylated on Cys-2 [109]. We generated palmitoylation deficient variants of APT1 and 2 as well as catalytically inactive versions (APT1 S119A and APT2 S121A). Overexpression of WT or mutant APT1 had no detectable effect on DHHC6 palmitoylation (Figure 5.3BC). In contrast, overexpression of WT, but not palmitoylation deficient, APT2 led to a significant decrease in DHHC6 palmitoylation. APT2 S121A had an intermediate effect, possibly due to the formation of heterodimers between mutant and endogenous APT2.

Altogether these observations show that DHHC6 palmitoylation is dynamic, in particular on Cys-328, and that depalmitoylation is mediated by APT2.

5.2.4 Palmitoylation of Cys-328 destabilizes DHHC6

We next determined whether palmitoylation has an influence on DHHC6 stability since palmitoylation was found to stabilize proteins such calnexin [39] and the death receptor Fas [128] and destabilize gp78 [121]. Stability was monitored using ³⁵S-Cys/Met metabolic pulse-chase. Myc-DHHC6 was transiently expressed in HeLa cells, which were submitted to a 20 min metabolic pulse followed by different times of chase before anti-myc immunoprecipitation, SDS-PAGE and auto-radiography. Decay of newly synthesized DHHC6 was biphasic, with 40% undergoing gradual degradation during the first 5 hrs, and the remaining 60% undergoing degradation at a greatly reduced rate (Figure 5.3A). Globally, degradation of 50 % of DHHC6 was observed at ≈16hrs. We repeated the experiments upon over expression of DHHC16, as well as upon silencing of APT2 expression, with the aim of increasing DHHC6 palmitoylation.

Both genetic manipulations led to a dramatic acceleration of DHHC6 decay, 50% being degraded by 2 and 3hrs in DHHC16 over expressing and APT2 silenced cells respectively (Figure 5.3A). To address the importance of the different palmitoylation sites, similar experiments were performed on single and double cysteine mutants (Figure 5.3A & S4). Remarkably, mutation of Cys-328 to alanine abolished the sensitivity to DHHC16 overexpression or APT2 silencing (Figure 5.3A & S4B), whereas the CAA and the CCA mutants showed the same sensitivity to over-palmitoylation as WT (Figure 5.3A & S4A). Together these observations indicate that palmitoylation of Cys-328 renders DHHC6 susceptible to degradation. This is mediated by ERAD since DHHC6 underwent ubiquitination in APT2 silenced cells (Figure 5.3B) and could be rescued by the proteasome inhibitor MG132 (Figure 5.3C).

5.2.5 Palmitoylation-dependent DHHC6 localization

Palmitoylation may affect the association of proteins with specific membranes or membrane domains [47]. We therefore investigated whether palmitoylation would affect DHHC6 localization. Whereas WT DHHC6 showed a typical ER staining with strong co-localization with BIP, a luminal ER chaperone, and BAP31, a transmembrane ER protein, DHHC6 AAA, although clearly present in the ER tended to accumulate in dot-like structures which stained positive for BAP31 but not BIP (Figure 5.4A).

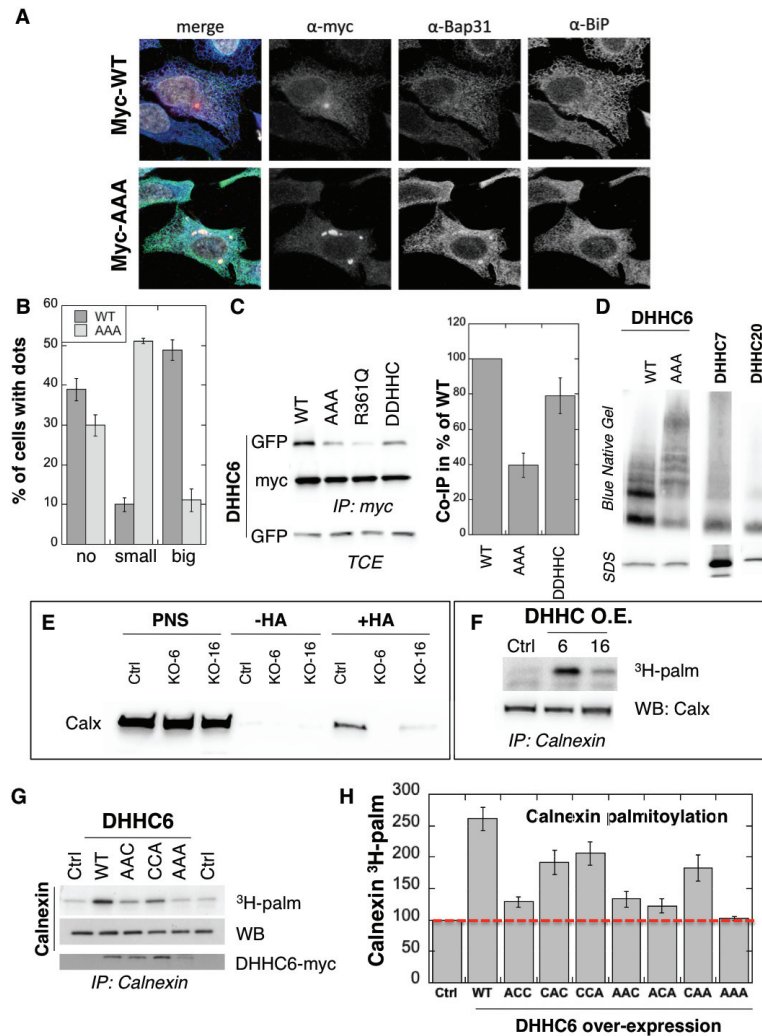


Figure 5.4. A. Immunofluorescence staining of HeLa cells transfected with PAT6-WT or PAT6-AAA mutant. B. HeLa cells were transfected with either PAT6-WT, PAT6-AAA or PAT6-RQ myc-tagged mutants and stained by immunofluorescence. The presence of dots (small or big) was reported for each cell and quantified. C. Co-immunoprecipitation of DHHC6 with mutant DHHC6. HeLa cells were transfected with plasmids encoding WT myc-DHHC6 and the indicated mutant DHHC6 GFP-tagged constructs for 24h. Proteins were extracted, a total cell extract was analyzed (TCE) and proteins were immunoprecipitated with myc antibodies and subjected to SDS-PAGE, then analyzed by immunoblotting with anti-myc or anti-GFP antibodies. Quantification of co-immunoprecipitation of WT DHHC6 with mutants DHHC6 constructs and of R361Q DHHC6 with mutants DHHC6 constructs. The calculated value of co-immunoprecipitation with WT DHHC6 or with R361Q DHHC6 mutant was set to 100% all DHHC6 mutants were expressed relative to this. N=4 for WT and N=3 for RQ361. D. DHHC6 complexes. HeLa cells were transfected with plasmids encoding WT myc-DHHC6 and the indicated mutant DHHC6 myc-tagged, or myc-tagged DHHC7 or DHHC20 constructs for 24h. Proteins were extracted, 40ug of a total cell extract was analyzed by SDS-PAGE or on blue native gels then analyzed by immunoblotting with anti-myc antibodies. E. Analysis of endogenous calnexin acylation in HAP WT control cells (Ctrl) versus HAP KO for DHHC6 (KO-6) or for DHHC16 (KO-16). Hap cell membranes were recovered by centrifugation and incubated with MMTS and then with hydroxylamine (+HA) or with TRIS (-HA) together with free thiol group binding beads. Eluted fractions were analyzed by immunoblotting with anti-calnexin antibodies. The input fraction (PNS) was loaded as 1/10 of this amount. F. Analysis of exogenous calnexin palmitoylation in HeLa cells overexpressing DHHC6 or DHHC16. HeLa cells were transfected with plasmids encoding WT myc-DHHC6 or WT myc-DHHC16 constructs for 24h. Cells were then metabolically labeled 2 hours at 37°C with 3H-palmitic acid. Proteins were extracted and immunoprecipitated with anti-calnexin antibodies, subjected to SDS-PAGE and analyzed by autoradiography (3H-palm) or by immunoblotting with anti-calnexin antibodies. G. Analysis of endogenous Calnexin palmitoylation in HeLa cells overexpressing DHHC6 mutants. HeLa cells were transfected with control plasmid (Ctrl) or plasmids encoding WT or mutants myc-DHHC6 constructs for 24h. Cells were then metabolically labeled 2 hours at 37°C with 3H-palmitic acid. Proteins were extracted and immunoprecipitated with anti-calnexin antibodies, subjected to SDS-PAGE and analyzed by autoradiography (3H-palm), quantified using the Typhoon Imager

or by immunoblotting with anti-calnexin or anti-myc antibodies. H. Quantification of ³H-palmitic acid incorporation into endogenous calnexin. The calculated value of ³H-palmitic acid incorporation into calnexin was set to 100% for control plasmid and all DHHC6 mutants were expressed relative to this. N=4.

Certain DHHC enzymes were reported to dimerize [129]. We therefore tested whether DHHC6 also dimerizes, and if so in a palmitoylation dependent manner. Immunoprecipitation experiments using DHHC6 constructs with two different tags showed that WT DHHC6 can associate with itself as well as with the Δ DHHC variant but not with the AAA mutant, or a single point mutant R361Q (Figure 5.4C) which also showed a dotted ER staining. AAA and R361Q however retained the ability to interact with themselves and with one another (Figure 5.4C). Blue native gel analysis confirmed that DHHC6 can associate into higher order structures, possibly dimers but also higher order complexes, as opposed to DHHC 7 and 20 which appeared largely monomeric (Figure 5.4D). The AAA mutant also formed higher order complexes but these were clearly different from those formed by WT protein (Figure 5.4D). Thus palmitoylation affects higher order assembly of DHHC6 and possibly thereby localization of the protein in the ER.

5.2.6 Palmitoylation dependent activity of DHHC6

Finally, we tested whether DHHC6 activity was modulated by palmitoylation. Calnexin palmitoylation was used as a readout for DHHC6 activity. Acyl-RAC analysis of DHHC16 KO cells indicated that palmitoylation of calnexin was significantly reduced (Figure 5.4E) and reversely that over-expression of DHHC16 led to an increase of calnexin palmitoylation (Figure 5.4F), together indicating that palmitoylation affects DHHC6 activity. Over-expression of WT DHHC6 itself led to 2.6 fold increase in calnexin palmitoylation (Figure 5.4GH), the background level being due to endogenous DHHC6 (Figure 5.4GH). Again this was palmitoylation dependent since all mutants had a lower activity than WT, the AAA mutant showing no activity at all

(Figure 5.4GH). Interestingly the presence of Cys-328 alone was sufficient to confer significant activity.

5.2.7 Model of the DHHC6 palmitoylation system

The above data shows that DHHC6 can be palmitoylated on 3 sites, that palmitoylation is dynamic and that it strongly influences localization, stability and function of the enzyme. The presence of 3 sites leads to the potential existence of 8 species: from fully unoccupied sites, termed C^{000} , to full occupancy, C^{111} (Figure 5.5A).

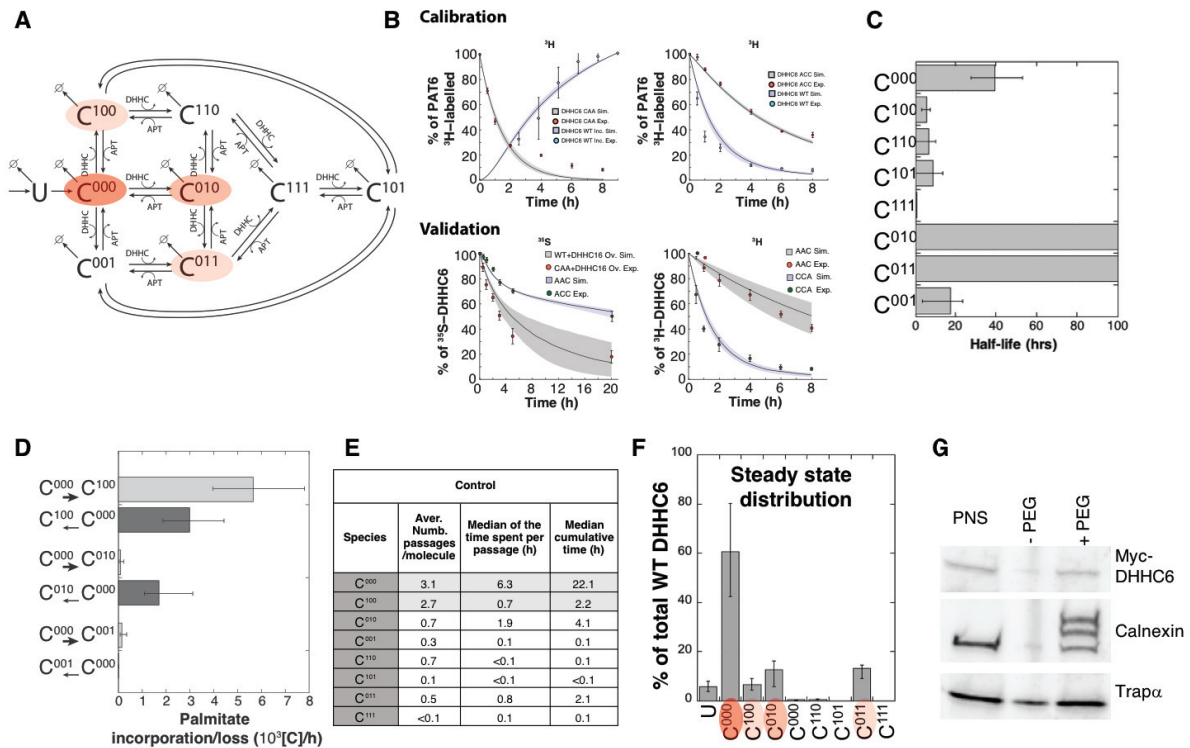


Figure 5.5. A. Network topology of the DHHC6 palmitoylation model. First we have a phase of synthesis of the unfolded peptide (U). The protein goes through a process of folding and membrane embedding, ending in the fully folded form of DHHC6 (C^{000}). The three sites can then be palmitoylated by DHHC16, the first palmitoylation can occur on each site, C^{100} C^{010} and C^{001} denote palmitoylation on the first, second or third site respectively. DHHC6 can then undergo another palmitoylation step acquiring two palmitates; C^{110} , C^{101} and C^{011} denote the double palmitoylated enzyme. From each of the three double palmitoylated states, DHHC6 can be modified one last time becoming fully palmitoylated (C^{111}). We consider that DHHC6 degradation can happen in each of this states. Red circles indicate the most abundant species in steady state. **B.** Part of the

calibration and validation sets used for parameter estimation. In the top graphs we show 3 curves that were used during parameter estimation with genetic algorithm to evaluate the goodness of fit of the parameters generated. The four curves in the bottom graphs were used to verify the accuracy of model predictions on experiments that were not used for parameter calibration. Circles represent experimental data while the solid lines are the output of the model after optimization. Since we have 152 different sets of optimal parameters, the shadows behind the lines represent the 1st and 3rd quartile of the 152 model outputs. The remaining part of the curves used for parameter estimation can be found in the supplementary material. C. Half-life of the different DHHC6 palmitoylation states estimated from the decay rate constants of the model after optimization. The half-life was calculated as: $\ln(2)/kd_i$. Where kd_i is the decay rate constant of the i th palmitoylation state. D. Palmitoylation and depalmitoylation fluxes for the three steps of single palmitoylation in steady state. Here measured the fluxes of palmitate incorporation and loss on the first, second and third site during the first palmitoylation event. Model fluxes in steady state for each reaction occurring in are available in the supplementary material. E. Results of single molecule tracking with stochastic simulations in control conditions. The table shows the average number of passage per molecule in each state of the model, along with the median and the cumulative median of the time spent in each state. This data were obtained from the analysis of 10000 stochastic simulations. The number reported here are the median of the model output of the 10000 stochastic simulations. F. Model prediction about DHHC6 WT steady state distribution in the different palmitoylation states. G. Stoichiometries of DHHC6 palmitoylation in Hela cells. Hela cells were transfected with plasmids encoding WT myc-DHHC6 constructs for 24h. Protein lysates were processed for the APEGS assay. PEG-5k were used for labeling of transfected myc-DHHC6 and endogenous protein (PEG+), PEG- lanes indicate the negative controls. The samples were analyzed by western blotting with anti-myc, anti,calnexin, anti-TRAPalpha antibodies.

To understand the dynamics of the inter-conversion between the 8 palmitoylation species, we developed a mathematical model (Figure 5.5A). Modelling was performed as an open system, including protein synthesis, and degradation of all species (Figure 5.5A). Synthesis of DHHC6 first leads to an unfolded species (U). Folding subsequently leads to the C^{000} species, which can undergo palmitoylation on any of the 3 sites. Single palmitoylated species can undergo a second and a third palmitoylation event. Each palmitoylation reaction is reversible. Since DHHC16 catalyses palmitoylation of all sites, a competition term between the sites was implemented in the enzymatic kinetics as we have previously done when modelling palmitoylation of the ER chaperone calnexin [39]. A similar competition term was introduced for depalmitoylation by APT2. The model also includes degradation rates for each species, with different first-order rate constants. The description of the rate expressions, the definition of the parameters, and the assumptions used in the development of the model are described in detail in the Supplementary Information.

To calibrate and subsequently validate the model, we performed the following set of kinetic experiments: 1) metabolic ^{35}S Cys/Met pulse-chase experiments, with pulses of either 20 min or 2 hrs, to monitor the stability of newly synthesized proteins, for WT and the different single, double and triple cysteine mutants; 2) ^3H palmitate incorporation into WT and cysteine mutants; 3) palmitate loss, monitored by ^3H

palmitate pulse-chase experiments, from WT and mutants. Some of these experiments were in addition performed upon over expression of DHHC16 or siRNA of APT2.

During calibration all these data were fitted simultaneously, transforming this process in a multi-objective optimization problem. These generally do not have a unique solution, but a population of sets of parameters that may fit different objectives differently, but are optimal for the entire set. We therefore employed a stochastic optimization method to generate a population of models consistent with the calibration experiments. From a population of 10'000 models, we selected 152 that based on an objective function fitted the experimental data most accurately (Suppl. Info.). The pool of selected models was subsequently used for the simulations and analyses. Importantly, all predictions were obtained by simulating each model independently. Outputs of all models were averaged and standard deviations with respect to the mean were calculated (Figure 5.5B). Importantly, the model reliably predicts the set of experiments that were used for the validation, indicating that it accurately captures the DHHC6 palmitoylation system (Figure 5.5B).

5.2.8 Dynamics of DHHC6 palmitoylation and effect on stability

The model was first used to estimate the half-lives of the difference species. C^{000} is predicted to have a half-life of approx. 40 hrs (Figure 5.5C and Table S5.1). The presence of palmitate on site 1 is predicted to strongly reduce the half-life of the species, irrespective of the occupancy of the two other sites, with a $t_{1/2}$ of 5 and 0.3 hrs for C^{100} and C^{111} respectively. In contrast, the presence of palmitate on site 2, in the absence of palmitate on site 1, is predicted to have a stabilizing effect with $t_{1/2}$ of C^{010} being over 100 hrs. Palmitoylation on site 3 is predicted to have a destabilizing effect with a $t_{1/2}$ of C^{001} of 18 hrs (Figure 5.5C). Thus, consistent with the experimental observations, the model indicates that palmitoylation strongly affects DHHC6 stability in a site dependent manner.

We next analysed the dynamics of the system. We first determined the palmitoylation and depalmitoylation fluxes. The major fluxes through the system are from C^{000} to C^{100} , backwards from C^{100} to C^{000} and to a lesser extent from C^{010} to C^{000} (Figure 5.5D). We then derived a stochastic formulation of the model, and performed 10'000 simulations to track single proteins in the system. DHHC6 molecules are predicted to spend most of their time, by far, in the C^{000} state (Figure 5.5E). Consistent with the flux analysis, each DHHC6 molecules on average was in the C^{100} state 2.7 times (Figure 5.5E), and remained in that state for 42 min. Seven out of 10 molecules also explored the C^{010} state for almost 2 hrs and more briefly the C^{011} state.

Consistent with this stochastic modelling, estimation of the steady state distribution of species under our experimental setting indicate that about 60% of the DHHC6 molecules in the cell are in the C^{000} state and only 25 % are palmitoylated, the most abundant species being C^{010} and C^{011} (Figure 5.5F). We tested this prediction experimentally by performing a PEGylation assay. This is a mass-tag labeling method that allows replacing palmitate with PEG through disruption of the thioester bond with hydroxylamine resulting in a mass change detectable by Western blotting [130]. As a control we analyzed calnexin, which has 2 palmitoylation sites and migrated as expected as 3 bands, corresponding to the non-, single and dual palmitoylated forms (Figure 5.5G). Only the non-palmitoylated form of DHHC6 was detected DHHC6, consistent with the prediction (Figure 5.5GF). Given the dynamic range and sensitivity of Western blots, a band with a 5 fold lower intensity than the C^{000} DHHC6 band would not be detectable.

Altogether, the model of DHHC6 palmitoylation predicts that under our experimental conditions, DHHC6 is mostly non-palmitoylated –which could be confirmed experimentally–, but that during their life cycle each molecule undergoes multiple rounds of palmitoylation-depalmitoylation on site 1, and for a subset of molecules also on site 2 and/or on site 3.

5.2.9 Importance of multiple palmitoylation sites for the regulation of DHHC6 activity

Figure 5.4H shows the relative activity of the various DHHC6 cysteine mutants. The most active variants are those that retain palmitoylation on site 1, suggesting that palmitoylation on site 1 most strongly promotes activity. Palmitoylation on site 1 however also has the strongest destabilizing effect (Figure 5.5C). This raises the question as to how cells could increase their DHHC6 activity, when needed, if increase in activity simultaneously leads to targeting of the enzyme to degradation. We hypothesized that sites 2 and 3 might be there to allow cells to increase the overall DHHC6 activity without significant loss of enzyme amounts.

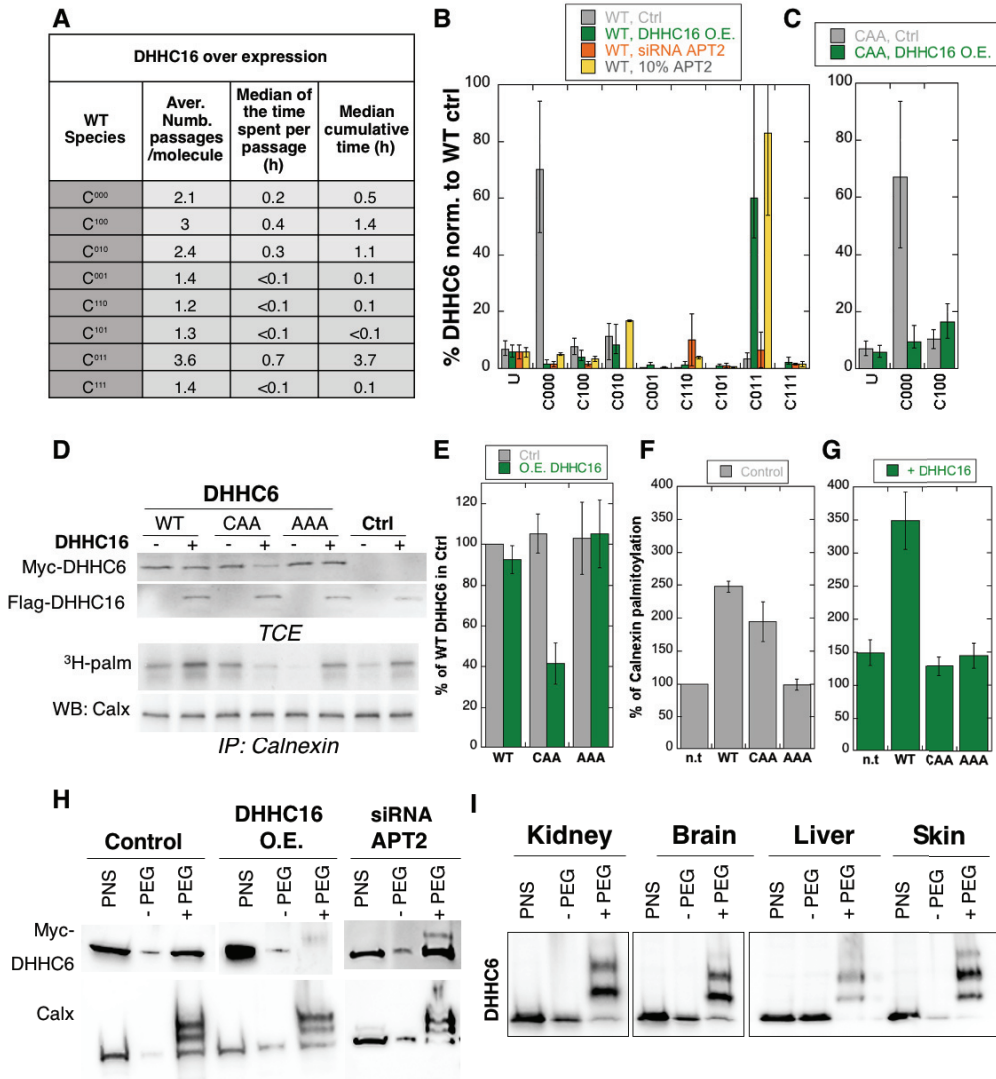


Figure 5.6. A. Results of single molecule tracking with stochastic simulations when DHHC16 is overexpressed. The table shows the average number of passage per molecule in each state of the model, along with the median and the cumulative median of the time spent in each state. As for control experiments this numbers were obtained by averaging the results of 10000 independent stochastic simulations. B. Steady state distribution of DHHC6 WT in the different palmitoylation states under different conditions (control, DHHC16 Overexpression, APT2 silencing and APT2 at 10% level of expression with respect to WT condition). All the data are scaled with respect to the total abundance of DHHC6 WT in normal condition. C. steady state distribution of DHHC6 CAA mutant under control condition or DHHC16 overexpression. All the data are scaled with respect to the total abundance of DHHC6 WT in normal condition. D. Calnexin palmitoylation level with DHHC6 and DHHC16 overexpressed. HeLa cells were transfected with control plasmid (Ctrl) or plasmids encoding WT or mutants myc-DHHC6 in the presence or not of flag-DHHC16 constructs for 24h. Cells were then metabolically labeled 2 hours at 37°C with 3H-palmitic acid. Proteins were extracted, TCE isolated, and immunoprecipitated with anti-calnexin antibodies, subjected to SDS-PAGE and analyzed by autoradiography (3H-palm), quantified using the Typhoon Imager or by immunoblotting with anti-calnexin, anti-myc or anti-flag antibodies. E. Quantification of DHHC6 levels in total cell extracts with or without DHHC16. The calculated value of WT DHHC6 in absence of overexpressed DHHC16 was set to 100%. All DHHC6 mutant with or without DHHC16 were expressed relative to this. N=3 F.G. Quantification of 3H-palmitic acid incorporation into endogenous

calnexin with DHHC6 in absence (F) or presence of DHHC16 (G). The calculated value of 3H-palmitic acid incorporation into calnexin was set to 100% for control plasmid (Ctrl) and all DHHC6 mutants were expressed relative to this. N=3. H. Stoichiometries of DHHC6 palmitoylation in HeLa cells with DHHC16 overexpressed or APT2 silenced. HeLa cells were transfected with plasmids encoding WT myc-DHHC6 with Flag-DHHC16 constructs for 24h or after 48h silencing APT2. Protein lysates were processed for the APEGS assay. PEG-5k were used for labeling of transfected myc-DHHC6 (PEG+), PEG- lanes indicate the negative controls. The samples were analyzed by western blotting with anti-myc. I. Adult mouse tissues were processed for the APEGS assay. PEG-5k were used for labeling of endogenous DHHC6 (PEG+), PEG- lanes indicate the negative controls. The samples were analyzed by western blotting with anti-DHHC6.

To test this, we first estimated the steady state species distribution of WT DHHC6 under conditions of hyperpalmitoylation, i.e. either when overexpressing DHHC16 or silencing APT2. Stochastic simulations indicate that overexpression of DHHC16 drastically changes the dynamics of the network (Figure 5.6A). Under these conditions, all molecules appear on average to explore all the palmitoylation states, with extremely short residence times in each. In terms of cumulative time, each DHHC6 molecules then spends 3.7 hrs in the C^{011} state, 1.4 hr in the C^{100} state, 1.1 hr in the C^{010} state and 30 mins or less in all the others (Figure 5.6A). These changes are also apparent from the steady state distribution, which shows that upon DHHC16 overexpression some 60% of the molecules are now in the most stable C^{011} state (Figure 5.6B). Interestingly, flux analysis indicates that the paths leading from C^{000} to C^{011} involve 4 steps (Figure 5.6C). Under DHHC16 overexpression conditions, C^{011} appears to be the hub of the system. During 10'000 simulations, 22'000 events of palmitoylation-depalmitoylation emanated from C^{011} whereas only 4'000 events occurred between C^{000} and C^{100} . In terms of overall protein level, DHHC16 overexpression was predicted to decrease the cellular amounts of DHHC6 by 15% (Table S5.2). APT2 siRNA also led to a species redistribution to the benefit of the dual and triple palmitoylated state (Figure 5.6B). The effect on protein abundance was however drastic, with only 30% expression when compared to control conditions (Table S5.2). Since the absence of APT2 is expected to freeze the dynamics Table S5.3, we estimated the situation under conditions where the APT2 activity would be reduced to 10% of control conditions. This mimicked the distribution obtained by DHHC16 overexpression (Figure 5.6B).

We next determined the steady state distribution of the CAA mutant, which can only adopt the C^{000} and C^{100} states. Under control conditions, the CAA species distribution was similar to that of WT, with the majority of the protein in the C^{000} state (Figure

5.6BC). DHHC16 overexpression however led to a drastic drop of the C⁰⁰⁰ population and an only mild increase in C¹⁰⁰ (Figure 5.6B), the total protein concentration decreasing by 70% (Table S5.2).

We next tested the prediction that the presence of sites 2 and 3 renders DHHC6 resistant, in terms of concentration, to fluctuations in the DHHC16 and APT2 activities. As seen in Figure 5.64B, SiRNA of APT2 indeed leads to a drastic decrease in DHHC6 expression. DHHC16 overexpression barely affected WT levels, as predicted, and did not change AAA levels, as expected (Figure 5.6DE). The levels of CAA however dropped by 60% (Figure 5.6DE). Consistently DHHC16 expression also led to an increase in WT DHHC6 activity (Figure 5.6F), as monitored by calnexin palmitoylation, but there was not increase in activity in cells expressing CAA vs. untransfected (Figure 5.6G). Note that the “background” calnexin palmitoylation level was increased due to the activation of endogenous DHHC6 by DHHC16 overexpression. PEGylation experiments confirmed that DHHC16 overexpression leads to an increase of the palmitoylated species (Figure 5.6H).

5.3 Discussion

S-palmitoylation consists in the covalent attachment of a palmitate molecule to a cysteine residue of a protein. This reversible post-translational modification has shown to possess an important role in regulation of the fate of different proteins, and as consequence on the regulation of different cellular functions. In fact, palmitoylation allow interaction of soluble protein with membranes, affects protein conformation, trafficking, stability and function.

While the pool of palmitoylated proteins is constantly increasing, little is known about the regulation of this modification. In this work we focused on one of the 23 human palmitoyltransferase, DHHC6, with the aim of investigating the mechanisms that regulate this post-translational modification. DHHC6 is a membrane protein resident

in the ER, where it controls the modification of a variety of proteins being on of the major responsible in regulation of their stability, localization trafficking and function. In this work we have shown that DHHC6 is itself palmitoylated by DHHC16 on three different cysteines (Cys-328, Cys-329, Cys-343). ³H labelling experiments show that all the three cysteines are palmitoylated, with Cys-328 being the most responsible for signal loss in case of mutation, suggesting that this cysteine is preferentially palmitoylated.

Since palmitoylation is a dynamic modification, we investigate this possibility on DHHC6 with ³H pulse-chase experiments. Results indicate a rapid turnover of DHHC6 palmitoylation, with a loss of 50% of the signal after only 1hr chase. Interestingly palmitate turnover slow down considerably in the absence of Cys-328, suggesting that palmitoylation of this site is more dynamic with respect to the others sites. The reaction of palmitate removal is catalysed by a class of enzymes called Acyl protein thioesterase, we therefore investigated their effects on DHHC6 depalmitoylation. While overexpression of APT1 didn't have any detectable effect on DHHC6 palmitoylation, overexpression of APT2 led to a significant decrease in DHHC6 palmitoylation. Altogether these observations show that DHHC6 palmitoylation is dynamic, in particular on Cys-328, and that depalmitoylation is mediated by APT2.

Next we investigated the effects of palmitoylation of each site on DHHC6. Since in a previous work we observed that palmitoylation stabilize the ER chaperone calnexin, we investigated if palmitoylation was affecting DHHC6 stability as well. Stability was monitored using ³⁵S-Cys/Met metabolic pulse chase, similarly to calnexin DHHC6 decay was biphasic, with fast degradation of 40% of the pool of tagged protein within the first 5 hours of chase while the remaining 60% underwent degradation at greatly reduce rate. Unexpectedly overexpression of DHHC16, or APT2 silencing, led to a considerable increase in DHHC6 decay, with 50% of the pool of protein degraded in the first 2-3hrs, suggesting that palmitoylation would actually destabilize the enzyme. To verify if each individual site had the same effect on stability, similar experiments were performed on single or double cysteine mutants. Remarkably, mutation of Cys-

328 to alanine abolished the sensitivity to DHHC16 overexpression or APT2 silencing, indicating that Cys-328 is responsible for rapid degradation of DHHC6. All together these observations indicate that DHHC6 stability is dependent on the palmitoylation pattern of the enzyme, with Cys-328 being responsible for a considerable drop of half-life. These results uncover a very interesting regulatory mechanism that allows the control of DHHC6 half-life depending on its palmitoylation state.

Since palmitoylation affects other proteins localization we investigated if this was the case for DHHC6 as well. While DHHC6 show ER staining and co-localization with BIP, a luminal ER chaperone, and the ER chaperone BAP31, a mutant version of DHHC6 which couldn't be palmitoylated show that the enzyme tends to accumulate on dot-like structures. These results show that palmitoylation strongly influence the localization of DHHC6 in the ER.

We next tested if the activity of DHHC6 was modulated by palmitoylation. Activity was measured by checking palmitate incorporation in calnexin. Remarkably palmitoylation of calnexin in DHHC16 KO cells was greatly reduced, and the opposite effect was observed upon DHHC16 overexpression. These results clearly indicate a strong correlation between DHHC6 activity and its palmitoylation state. Overexpression of DHHC6 and mutants shows that the presence of the three sites is required for proper activity of the enzyme, in fact all the mutants had a lower activity than WT, with AAA mutant showing no activity at all. Notably the CAA mutant shows an amount of activity comparable to WT, suggesting that this site is the main responsible in regulation of DHHC6 activity.

In the second part of this work, results from experimental analysis were used to build a kinetic model of DHHC6 palmitoylation. The model was used to study the properties and the dynamics of inter-conversion between the 8 palmitoylation states of DHHC6. We first investigated the half-life of each palmitoylation pattern, to get more insight on the effect of each palmitoylation site on the stability of DHHC6. While the non palmitoylated form of DHHC6 is very stable, with an estimate of 40hrs, the presence of palmitate on site 1 lead to a huge drop in half-life, in agreement with

what was found experimentally. In addition, the model suggest that the destabilization effect of site 1 is irrespective of the occupancy of the two other sites, since all the palmitoylation pattern in which site 1 is modified have an half-life in between 5 and 0.3 hrs. In contrast, palmitoylation of site 2 and 3 in absence of modification of site 1 have a huge stabilizing effect, increasing the half-life of the protein to more than 100 hours. Palmitoylation on site 3 alone show a moderate destabilizing effect, decreasing the half-life of the protein to 18hrs. These results show that palmitoylation of DHHC6 can positively or negatively regulate the stability of the enzyme based on the palmitoylation pattern, demonstrating the primary role of palmitoylation in regulating the fate of the protein.

Analysis of the model dynamics show that there are consistent palmitoylation/depalmitoylation fluxes between C^{000} and C^{100} , and to a lesser extent between C^{000} and C^{010} . In agreement with the experiments the model suggested consistent cycles of palmitoylation on site 1, highlighting the role of this site as primary way to regulate stability and activity of the enzymes.

To better investigate the dynamics between the different palmitoylation states we developed a stochastic version of the model that allowed us to follow single DHHC6 proteins through the model. 10000 simulations of the model steady state allowed us to estimate some dynamic properties characteristic of DHHC6 proteins. Unexpectedly the enzyme shows to spend most part of its time in the unpalmitoylated form, which is also the inactive one. Many palmitoylation cycles were observed between C^{000} and C^{100} , confirming the results of the previous analysis. In normal conditions we measured that on average a single molecule of DHHC6 cycles 3 times between C^{000} and C^{100} , with a short resident time on C^{100} of less than 1 hour, in order to avoid depletion of the enzyme. DHHC6 also populated the C^{010} state, but less frequently than C^{100} and for shorter periods. Analysis of states population in normal condition reveal that the enzyme is mainly palmitoylated on site 1, with very low time spent in other single palmitoylated states. Little to no events of double palmitoylation are observed in normal conditions. Results to suggest that in normal conditions

palmitoylation happens almost exclusively on site 1, which is the major contributor in DHHC6 activity and stability.

Stochastic simulations were repeated also in condition of DHHC16 overexpression or APT2 silencing, causing overpalmitoylation of DHHC6. As consequence the enzyme undergoes dramatic changes in both its palmitoylation distribution and dynamics. The bulk of the protein shifted from a non palmitoylated state to C^{011} , one of the most stable state, replacing C^{000} as hub of the network dynamics. While DHHC6 in normal condition cycle mainly between C^{000} and C^{100} , now the enzyme cycle most between C^{011} and C^{001} , but an increased number of cycles is observed also for all the other states. Increasing the palmitate incorporation rates of DHHC6 by DHHC16 overexpression or APT2 silencing decreased the total cellular amount of DHHC6 by 15%. Notably some degree of APT2 activity is needed to avoid complete depletion of the enzyme, which in case of total APT2 silencing lose 70% of its pool. We estimated that an amount of APT2 activity of at least 10% of control conditions is needed to contain the drop in concentration caused by hyperpalmitoylation of DHHC6.

The same analysis was repeated on the CAA mutant of DHHC6, to investigate the contribution of site 2 and 3 on the effect of DHHC6 concentration. Remarkably the model predicts that since CAA can't reorganize itself in the C^{011} state DHHC16 overexpression led to a 70% drop in in steady state concentration. All together this results suggest the interesting possibility that sites 2 and 3 might be there to allow cells to increase the overall DHHC6 activity without significant loss of enzyme amounts. Activity of DHHC6 depend mainly on the amount of palmitoylation, but since palmitoylation of site 1 decrease the stability of the enzyme, the enzyme can reorganize its distribution into more stable palmitoylation states, which allow to increase the palmitoylation level of the enzyme, and therefore its activity, without decreasing significantly its stability. The analysis performed in this work have demonstrated the huge impact that palmitoylation have on the regulation of DHHC6 activity and concentration.

5.4 Material and methods

5.4.1.1 Cell lines

Hela cells (ATCC) were grown in complete modified Eagle's medium (MEM, Sigma) at 37°C supplemented with 10% fetal bovine serum (FBS), 2mM L-Glutamine, penicillin and streptomycin. For the DHHC6 knockdown cell lines, Hela cells were transfected with shRNA against DHHC6 gene (target sequence in 3'UTR: 5'-CCTAGTGCCATGATTTAAA-3') or with shRNA control against firefly luciferase gene (target sequence : 5'-CGTACGCGGAATACTTCGA-3'). The transfected cells were selected by treatment with 3µg/ml puromycin. HAPI Wild type WT and knockout cell lines were purchased from Horizon Genomics (Vienna, Austria). The DHHC6 clone (13474-01) contains a 5bp deletion in exon 2 (NM_022494) and the DHHC16 clone (36523-06) contains a 2bp insertion in exon 2 (NM_198043). HAPI cells were grown in complete Dulbecco's MEM (DMEM, Sigma) at 37°C supplemented with 10% fetal bovine serum (FBS), 2mM L-Glutamine, penicillin and streptomycin.

5.4.1.2 Antibodies and Reagents

The following primary antibodies are used: Mouse anti-Actin (Millipore, MAB 1510), Mouse anti-Myc 9E10 (Covance, MMs-150R), Mouse anti-Ubiquitin (Santa Cruz, sc-8017), Mouse anti-GFP (Roche, 11814460001), Rabbit anti-DHHC6 (Sigma, SAB304457), Mouse anti-Transferrin Receptor (Thermo Scientific, 136800), Mouse anti-Flag M2 (Sigma, F3165), Rabbit anti-ANTXR1 (Sigma, SAB2501028), Rabbit anti-TRAPalpha (Abcam, ab133238), Rabbit anti-Flotillin1 were produced in our laboratory, Mouse anti-CLIMP63 (Enzo, ALX-804-604), Mouse anti-Calnexin (MAB3126), Rabbit anti-GP78 AMFR (Abnova, PAB1684), Rabbit anti-IP3R (Cell signaling, 85685). The following beads were used for immunoprecipitation: Protein G Sepharose 4 Fast flow (GE Healthcare, 17-0618-01), anti-Myc affinity gel (Thermo Scientific, 20169), anti-Flag affinity gel EZview M2 (Sigma, F2426). Drugs were used as follows: Bafilomycin A1 at

100 nM (Sigma, B1793), MG132 at 10 uM (Sigma, C2211), Hydroxylamine at 0.5 M (Sigma, 55459), mPEG-5k at 20mM (Sigma, 63187), N-ethylmaleimide NEM at 20mM (Thermo Scientific, 23030), Tris-2- carboxyethyl-phosphine hydrochloride TCEP at 10 mM (Thermo Scientific, 23225), Methyl methanethiosulfonate MMTS at 1.5% (Sigma, 208795), Puromycin at 3ug/ml (Sigma, P9620).

5.4.1.3 Transfection and siRNA experiments:

Human Myc -DHHC6, Myc-DHHC6- C328A (ACC), Myc-DHHC6-C329A (CAC), Myc-DHHC6-C343A (CCA), Myc-DHHC6-C328A-C329A (AAC), Myc-DHHC6-C326A-C343A (ACA), Myc-DHHC6-C329A-C343A (CAA), Myc-DHHC6-C328A-C329A-C343A, Myc-DHHC6-R361Q, Myc-DHHC6-R361A, Myc-DHHC6-Del-K410-K411-N412-R413, Myc-DHHC6-Del-D126-H127-H128-C129 were cloned in pcDNA3. Human GFP-DHHC6 and all cyteine mutants mentioned above were also cloned in peGFP. Human FLAG-DHHC16 was cloned in pCE-puro-3xFLAG, human Myc-DHHC16 was cloned in pCE puro-his-myc. All other human Myc-DHHCs were cloned in pcDNA3.1 (provided by the Fukata lab). mCitrine fusions of APTs were inserted into pcDNA3.1-N1 (provided by Bastiaens lab, [108]). mCitrine APT1-S119A, mCitrine APT2-S121A, mCitrine-APT1-C2S, mCitrine-APT2-C2S were cloned in pcDNA3.1-N1. For control transfection, we used an empty pcDNA3 plasmid. Plasmids were transfected into HeLa cells for 24h (2ug cDNA/9.6cm²) plate using Fugene (Promega).

For gene silencing, Hela cells were transfected for 72 h with 100pmol/9.2cm² dish of siRNA using interferin (Polyplus) transfection reagent. As control siRNA we used the following target sequence of the viral glycoprotein VSV-G: ATTGAACAAACGAAACAAGGA. siRNA against human genes were purchased from Qiagen (DHHC6 target sequences : 1-GAGGTTTACGATACTGGTTAT, 2-TAGAAGGTGTTTCAAGAATAA, DHHC16 target sequences: 1-CTCGGGTGCTCTTACCTTCTA, 2-TAGCATCGAAAGGCACATCAA; APT1 target sequence: AACAACTTATGGGTAATAAAA; APT2 target sequence:

AAGCTGCTGCCTCCTGTCTAA, all other DHHC target sequences were previously tested [58]).

5.4.1.4 Real-time PCR

For HeLa cells, RNA was extracted from a six-well dish using the RNeasy kit (Qiagen). 1 mg of the total RNA extracted was used for the reverse transcription using random hexamers and superscript II (Thermo Scientific). A 1:40 dilution of the cDNA was used to perform the real-time PCR using SyBr green reagent (Roche). mRNA levels were normalized using three housekeeping genes: TATA-binding protein, β -microglobulin and β -glucuronidase. Total RNA of different mouse tissues were extracted using the RNeasy kit (Qiagen) after solubilization with TissueLyser II (Qiagen).

5.4.1.5 Radiolabelling experiments

For the ^{35}S -metabolic labeling, the cells were starved in DMEM HG devoid of Cys/Met for 30 minutes at 37°C, pulsed with the same medium supplemented with 140 μCi of ^{35}S Cys/Met (American Radiolabeled Chemicals, Inc.) for the indicated time, washed and incubated in DMEM complete medium for the indicated time of chase [42] before immunoprecipitation. To detect palmitoylation, HeLa cells were transfected or not with different constructs, incubated for 2 hours in IM (Glasgow minimal essential medium buffered with 10mM HEPES, pH 7.4) with 200 $\mu\text{Ci}/\text{ml}$ 3H palmitic acid (9,10- $^3\text{H}(\text{N})$) (American Radiolabeled Chemicals, Inc.). The cells were washed, incubated in DMEM complete medium for the indicated time of chase, or directly lysed for immunoprecipitation with the indicated antibodies.

For all radiolabelling experiments, after immunoprecipitation, washes beads were incubated for 5 min at 90°C in reducing sample buffer prior to 4-12% gradient SDS-PAGE. After SDS-PAGE, the gel are incubated in a fixative solution (25% isopropanol, 65% H₂O, 10% acetic acid), followed by a 30 min incubation with signal enhancer

Amplify NAMP100 (GE Healthcare). The radiolabeled products were revealed using Typhoon phosphoimager and quantified using the Typhoon Imager (ImageQuanTool, GE Healthcare).

5.4.1.6 Immunoprecipitation

For immunoprecipitation, cells were washed 3 times PBS, lysed 30min at 4°C in the following Buffer (0.5% Nonidet P-40, 500 mM Tris pH 7.4, 20 mM EDTA, 10 mM NaF, 2 mM benzamidin and protease inhibitor cocktail (Roche)), and centrifuged 3min at 5000 rpm. Supernatants were subjected to preclearing with G sepharose beads prior immunoprecipitation reaction. Supernatants were incubated overnight with the appropriate antibodies and G Sepharose beads.

5.4.1.7 Post nuclear supernatants and ACYL-RAC

HeLa cells were harvested, washed with PBS, and homogenized by passage through a 22G injection needle in HB (HB: 2.9 mM imidazole and 250 mM sucrose, pH 7.4) containing a mini tablet protease inhibitor mixture (Roche). After centrifugation, the supernatant was collected as PNS (Post Nuclear Supernatant).

Protein S-palmitoylation was assessed by the Acyl-RAC assay as described [131], with some modifications. A fraction of the PNS was saved as the input. HeLa PNS were lysed in buffer (0.5% Triton-X100, 25 mM HEPES, 25 mM NaCl, 1 mM EDTA, pH 7.4 and protease inhibitor cocktail). In order to block free SH groups with S-methyl methanethiosulfonate (MMTS), 200 µl of blocking buffer (100 mM HEPES, 1 mM EDTA, 87.5 mM SDS and 1.5% (v/v) MMTS) was added to cell lysate and incubated for 4 h at 40 °C. Subsequently, 3 volumes of ice-cold 100% acetone was added to the blocking protein mixture and incubated for 20 minutes at -20 °C and then centrifuged at 5,000 × g for 10 minutes at 4 °C to pellet precipitated proteins. The pellet was washed five times in 1 ml of 70% (v/v) acetone and resuspended in buffer (100 mM

HEPES, 1 mM EDTA, 35 mM SDS). For treatment with hydroxylamine (HA) and capture by Thiopropyl Sepharose® beads, 2 M hydroxylamine was added together with the beads (previously activated for 15 min with water) to a final concentration of 0.5 M hydroxylamine and 10% (w/v) beads. As a negative control, 2 M Tris was used instead of hydroxylamine. These samples were then incubated overnight at room temperature on a rotate wheel. The beads were washed, the proteins were eluted from the beads by incubations in 40 µl SDS sample buffer with beta-mercapto-ethanol for 5 minutes at 95 °C. Finally, samples were submitted to SDS-PAGE and analyzed by immunoblotting.

5.4.1.8 APEGS assay

The level of protein S-palmitoylation was assessed as described [115], with minor modifications. Hela cells were lysed with the following buffer (4% SDS, 5 mM EDTA, in PBS with complete inhibitor (Roche)). After centrifugation at $100,000 \times g$ for 15 min, supernatant proteins were reduced with 25 mM TCEP for 1 h at 55°C or at room temperature (RT), and free cysteine residues were alkylated with 20 mM NEM for 3 h at RT to be blocked. After chloroform/methanol precipitation, resuspended proteins in PBS with 4% SDS and 5 mM EDTA were incubated in buffer (1% SDS, 5 mM EDTA, 1 M NH₂OH, pH 7.0) for 1 h at 37°C to cleave palmitoylation thioester bonds. As a negative control, 1 M Tris-HCl, pH7.0, was used. After precipitation, resuspended proteins in PBS with 4% SDS were PEGylated with 20 mM mPEGs for 1 h at RT to label newly exposed cysteinyl thiols. As a negative control, 20 mM NEM was used instead of mPEG (-PEG). After precipitation, proteins were resuspended with SDS-sample buffer and boiled at 95°C for 5 min. Protein concentration was measured by BCA protein assay.

5.4.1.9 BLUE Native PAGE

The PNS of HeLa cells were extracted. The proteins were lysed in 1% Digitonin and passed through a 26g needle, incubated on ice 30 minutes, spin 16 000g for 30 minutes and run following the manufacturer instructions on the Novex Native PAGE Bis-tris gel system (ThermoFisher).

5.4.1.10 Immunofluorescence staining and fluorescence microscope

HeLa cells were seeded on 12mm glass coverslips (Marienfeld GmbH, Germany) 24hr prior to transfection. PAT6-myc plasmids were transfected using Fugene (Promega,USA) for 48hrs. Cells were then fixed using 3% paraformaldehyde for 20min at 37°C, quenched 10min with 50mM NH₄Cl at RT and permeabilized with 0.1% TX100 for 5min at RT and finally blocked overnight with 0.5% BSA in PBS. Cells were washed 3x with PBS in between all the steps. Cells were then stained with anti-myc and anti-BiP antibodies for 30min at RT, washed and incubated again 30min with their corresponding fluorescent secondary antibodies. Finally cells were mounted on glass slide using mowiol. Imaging was performed using a confocal microscope (LSM710, Zeiss, Germany) with a 63x oil immersion objective (NA 1.4).

5.4.1.11 Core model of DHHC6 palmitoylation

The DHHC6 palmitoylation model (Figure 5.5A) was designed following the approach elaborated previously for calnexin [39], a substrate of DHHC6. The core model is based on a previously described protein phosphorylation model used in [132]. The set of reaction described by Goldbeter was used to model a single palmitoylation event. Multiple palmitoylation events were modeled replicating this subunit for each reaction.

In Goldbeter work the model was mathematically described using mass action terms. Here, due to the presence of multiple modification events, which require the

definition of a consistent number of parameters, we described model reactions using what is called “total quasi-steady state approximation ” (tQSSA) [133,134]. With respect to the standard quasi-steady state approximation (QSSA), tQSSA is valid also when the enzyme substrate concentrations are comparable [135]. Since DHHC6 is palmitoylated by another palmitoyltransferases, namely DHHC16, the use of tQSSA is justified by the fact that different proteome studies suggest that DHHCs enzymes have similar concentrations. The step-by-step application of the tQSSA approximation to a palmitoylation model is described in [39].

The model can be divided in two parts; we first have a phase of synthesis of the unfolded peptide of DHHC6 (U in Figure 5.5A). The unfolded peptide goes through a process of folding and membrane embedding, ending in the fully folded form of DHHC6 (C^{000}), which shows the three sites available for palmitoylation on the cytosolic side of the ER membrane. The three sites can then be palmitoylated by DHHC16, the first palmitoylation can occur on each of the sites, C^{100} , C^{010} and C^{001} . These species can then undergo a second palmitoylation step, C^{110} , C^{101} and C^{011} , which can then undergo a third leading to C^{111} . Since palmitoylation is reversible the palmitate can be removed from each of the sites.

DHHC6 model is based on the following assumption:

1. DHHC6 can be degraded in each of its states.
2. DHHC6 is present in similar concentrations with respect to the modifying enzyme DHHC16 [127,136–139].
3. Acyl protein thioesterases (APTs) are more abundant than DHHC6 [127,136–139]. The three palmitoylation sites may have different affinities with respect to the palmitoylation/depalmitoylation enzymes, so we adopted separated K_{ms} and V_{maxs} for the different sites.
4. All the palmitoylation steps are reversible. APT catalyses the depalmitoylation steps.
5. Palmitate was considered to be available in excess inside the cell.

6. A detailed account of all reactions and differential equations in the model is given in Table S5.4-7.

5.4.1.12 Parameterization of the model

Since no kinetic data on DHHC6 palmitoylation is available, all the parameters of the model were estimated using a genetic algorithm. For the parameter optimization, multiple datasets coming from experimental results were considered. Time course labeling experiments were performed in order to characterize the dynamics of DHHC6 synthesis/degradation and incorporation/loss of palmitate.

The types of data used, the description of the genetic algorithm and the step-by-step application of the optimization algorithm used to find values for the parameters are described in detail in [39]. In this paper we use the exact same procedure.

For the parameter estimation we used a calibration set that correspond to the data visible in (Figure 5.5B and Figure S5.11). The remaining part of the data (Figure 5.5B and Figure S5.12) was used to validate the output of the model and to verify the accuracy of its prediction capabilities.

Because of the presence of multiple objectives there does not exist a single solution that simultaneously optimizes each objective, so the algorithm provides as output a local Pareto set of solutions, which are equally optimal with respect to the fitness function we defined. From the Pareto set provided by the GA and in order to be more accurate and to reduce the variability in the output of the model, we selected the set of parameters that best fitted the calibration data. The procedure for the selection of the subset is described in [39], 152 different sets of parameter were selected at the end of the optimization.

The results of the optimization can be found in Table S5.5

5.4.1.13 Simulating the labeling experiments

Since the goodness of each set of parameters is evaluated computing the distance between the output of the model and different experiments present in the calibration dataset, we made use of a previously established method to reproduce the different type of experiment *in-silico* [39].

5.4.1.14 Stochastic simulations

In order to measure the average palmitoylation time of DHHC6, we made use of a previously established method to perform *in-silico* single molecule tracking [39].

5.4.1.15 Conversion of deterministic parameters to stochastic:

Parameter conversion from deterministic to stochastic model is needed to perform stochastic simulations [39]. This transformation involves a change of units, from concentration to number of molecules. A single assumption was added to those previously defines [39] namely that the number of DHHC6 molecules per HeLa cell is around 1600 molecules [127,136–139] Table S5.7 shows the parameters obtained through the conversion. The model design is shown through the stoichiometry matrix and the propensity function in Table S5.8-10.

5.5 Appendix B

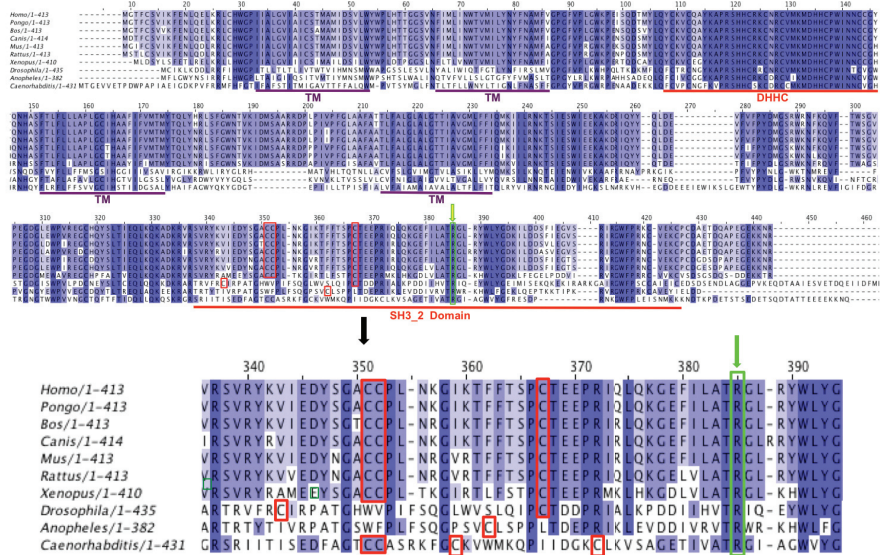


Figure S5.7. Alignment of DHHC6 sequences from different species. Identical amino acids are highlighted by dark blue. Palmitoylated cysteines are highlighted in red boxes. R361 is highlighted in yellow box. SH3-2 like domain and DHHC domains are underlined in red.

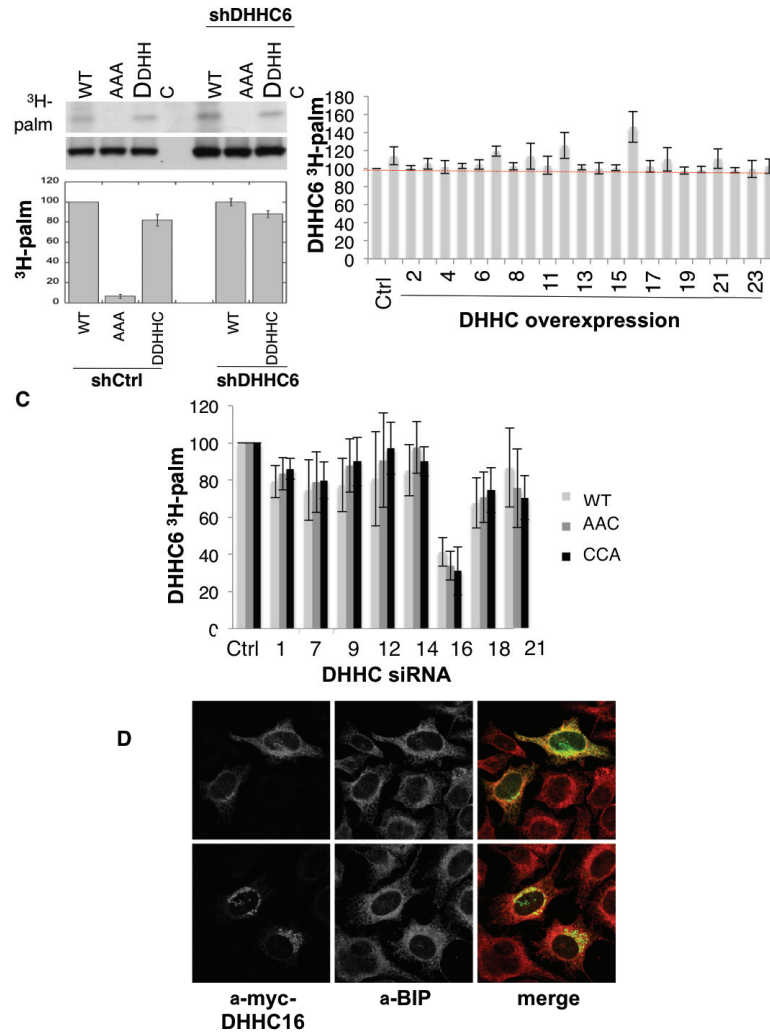


Figure S5.8. A. DHHC6 is not autopalmitoylated. HeLa cells silenced with control lentiviruses or with shDHHC6 lentiviruses were transfected with plasmids encoding WT myc-DHHC6 or mutant myc-DHHC6. Cells were then metabolically labeled 2 hours at 37°C with 3H-palmitic acid. Proteins were extracted and immunoprecipitated with anti-myc antibodies, subjected to SDS-PAGE and analyzed by autoradiography (3H-palm), quantified using the Typhoon Imager or by immunoblotting with anti-myc antibodies. 3H-palmitic acid incorporation into DHHC6 values were calculated and were set to 100% for WT DHHC6 constructs and all DHHC6 mutants were expressed relative to this. N=4. B. Palmitoylation of WT DHHC6 after human DHHC overexpression. HeLa cells were transfected with indicated DHHC constructs for 24h with WT myc-tagged DHHC6 construct. Cells were then metabolically labeled 2 hours at 37°C with 3H-palmitic acid. Proteins were extracted, immunoprecipitated with myc antibodies and subjected to SDS-PAGE and analyzed by autoradiography, quantified using the Typhoon Imager or by immunoblotting with myc antibodies. 3H-palmitic acid incorporation into WT DHHC6 constructs were quantified and normalized to protein expression level. The calculated value of 3H-palmitic acid incorporation into WT DHHC6 was set to 100% for a non relevant plasmid (Ctrl) and all DHHC were expressed relative to this. N=5. C. All DHHC6 cysteine are palmitoylated by DHHC16. HeLa cells were transfected with siRNA silencing indicated DHHC for 72h and with WT and cysteine mutants myc-tagged DHHC6 construct for the last 24h. Cells were then metabolically labeled 2 hours at 37°C with 3H-palmitic acid. Proteins were extracted, immunoprecipitated with myc antibodies and subjected to SDS-PAGE and analyzed by autoradiography, quantified using the Typhoon Imager or by immunoblotting with myc antibodies. 3H-palmitic acid incorporation into WT and mutants DHHC6 constructs were quantified and normalized to protein expression level. The calculated value of 3H-palmitic acid incorporation into WT or mutants DHHC6 was set to 100% for a non relevant siRNA (Ctrl) and all siRNA were expressed relative to this. N=3. D. Immunofluorescence staining of HeLa cells transfected with PAT16-WT-myc and imaged at high (top lane) or lower expression levels (bottom lane).

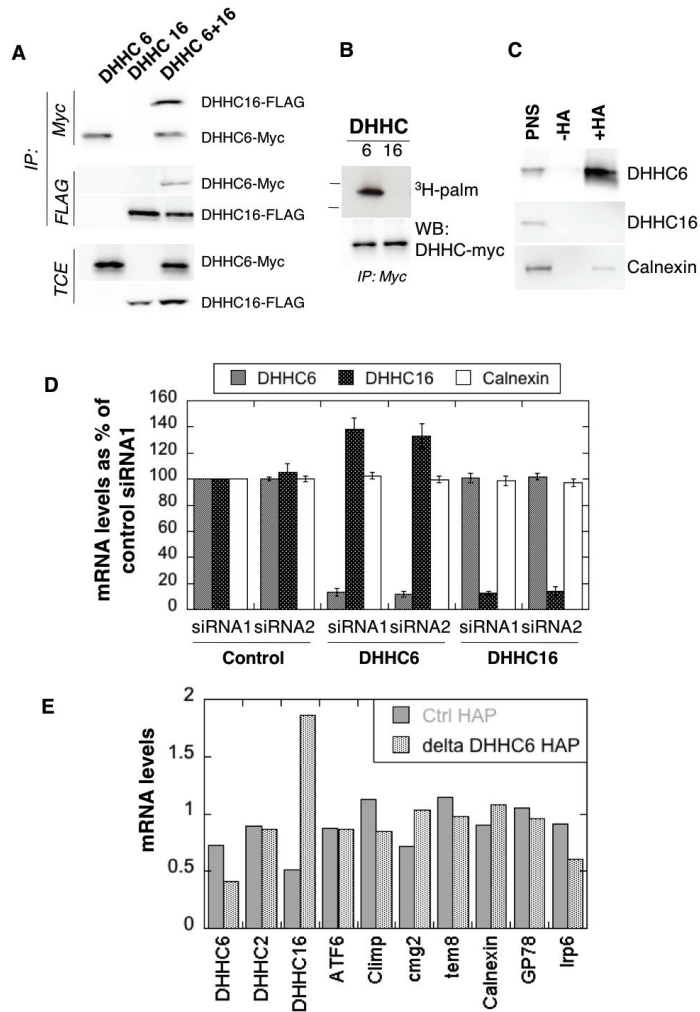


Figure S5.9. A. Co-immunoprecipitation of DHHC6 with DHHC16. HeLa cells were transfected with plasmids encoding WT myc-DHHC6 and flag-tagged DHHC16 constructs for 24h. Proteins were extracted, a total cell extract was analyzed (TCE) and proteins were immunoprecipitated with myc or flag antibodies and subjected to SDS-PAGE, then analyzed by immunoblotting with anti-myc or anti-flag antibodies. B. DHHC6 is palmitoylated but not DHHC16. HeLa cells were transfected with WT myc-tagged DHHC6 or flag-tagged DHHC16 construct for 24h. Cells were then metabolically labeled 2 hours at 37°C with 3H-palmitic acid. Proteins were extracted, immunoprecipitated with myc or flag antibodies and subjected to SDS-PAGE and analyzed by autoradiography, quantified using the Typhoon Imager or by immunoblotting with myc or flag antibodies. C. Analysis of protein acylation in HeLa cells (Ctrl). HeLa were transfected 24h with WT myc-DHHC6 or WT flag-DHHC16 constructs. Cell membranes were recovered by centrifugation and incubated with MMTS and then with hydroxylamine (+HA) or with TRIS (-HA) together with free thiol group binding beads. Eluted fractions were analyzed by immunoblotting with the indicated antibodies. The input fraction (PNS) was loaded as 1/10 this amount. D. Silencing of DHHC6 and DHHC16 is efficient. HeLa cells were transfected for 72h with either 2 control siRNA or different DHHC siRNA. DHHC6, DHHC16, Calnexin RNA levels were analyzed by quantitative RT-PCR. The histogram shows that silencing was efficient for all DHHC enzymes. N=4. E. DHHC16 RNA is increased in DHHC6 KO cells. Annotated RNA levels were analyzed by quantitative RT-PCR in HAP1 cells control or KO for DHHC6.

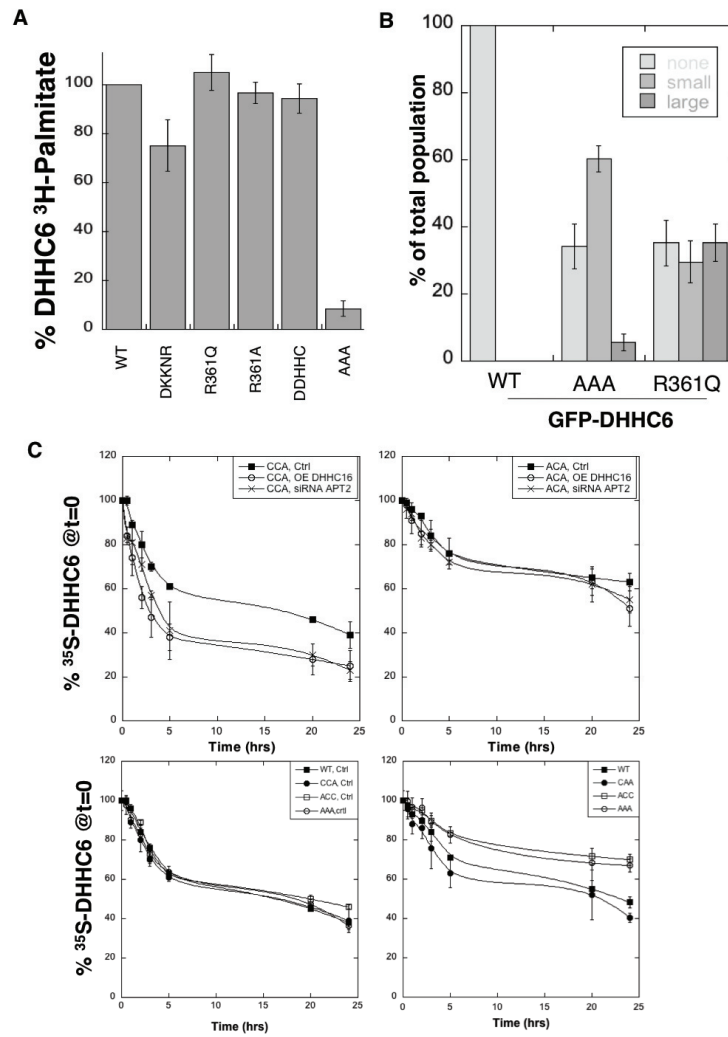


Figure 5.10. A. Level of mutant DHHC6 palmitoylation. HeLa cells were transfected with plasmids encoding WT or the indicated mutant MYC-tagged DHHC6 constructs for 24h. Cells were then metabolically labeled 2 hours at 37°C with 3H-palmitic acid. Proteins were extracted, immunoprecipitated with myc antibodies and subjected to SDS-PAGE and analyzed by autoradiography and quantified using the Typhoon Imager. Quantified values were normalized to protein expression level. The calculated value of 3H-palmitic acid incorporation into WT DHHC6 was set to 100% and all mutants were expressed relative to this. N=5. B. SH3 Mutant DHHC6 expression induces dots formation. HeLa cells were transfected with plasmids encoding WT or the indicated mutant MYC-tagged DHHC6 constructs for 24h. C. Degradation kinetics of DHHC6. HeLa cells were transfected with plasmids encoding WT flag-DHHC16, or WT myc-DHHC6 or cysteine mutants myc-DHHC6 constructs for 24h after 48h transfection with siRNA APT2 or with control siRNA. HeLa cells were incubated 20min pulse with ³⁵S-methionin/cysteine at 37°C, washed and further incubated for different times at 37°C in complete medium. DHHC6 were immunoprecipitated and subjected to SDS-PAGE and analyzed by autoradiography, quantified using the Typhoon Imager, and western blotting with anti-myc antibodies. ³⁵S-methionin/cysteine incorporation into different DHHC6 constructs were quantified for each times, normalized to protein expression level. The calculated value of ³⁵S-methionin/cysteine incorporation into DHHC6 was set to 100% for t=0 after the 20min pulse and all different times of chase with complete medium were expressed relative to this. N=3.

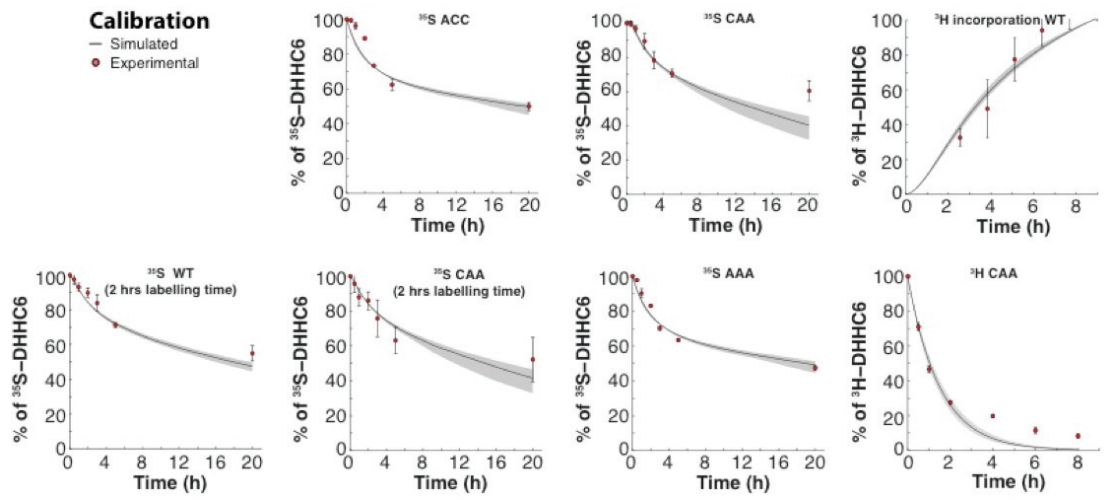


Figure S5.11. Calibration. Calibration data Results of the GA optimization are plotted on top of the experimental data used as objectives.

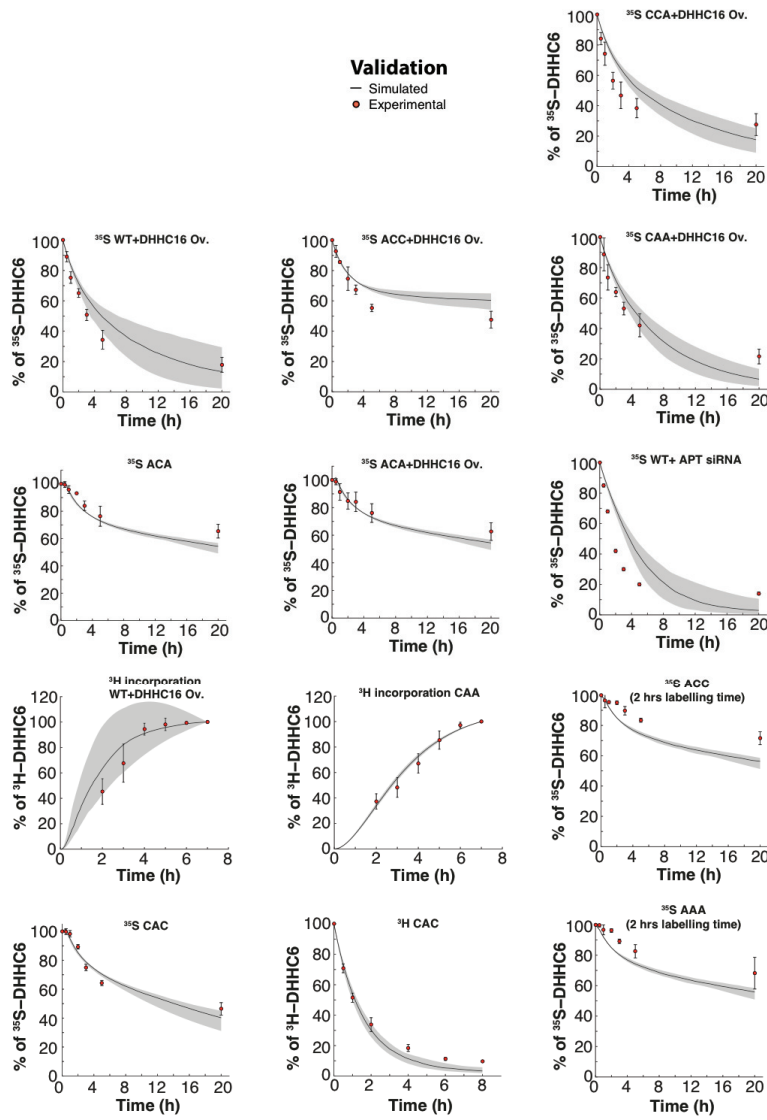


Figure S5.12. Validation. Results of the GA optimization are plotted on top of the experimental data used as validation.

Table S5.1 Half-life of DHHC6 in different palmitoylation states. The half-life was estimated from the decay rate constant obtained through parameter estimation. The half-life is calculate as: $\ln(2)/kd$.

Species	Half-Life (h)	Range (1 st and 3rd quartiles)
C^{000}	39.5	25.71-51.27
C^{100}	5.22	3.47-7.63
C^{010}	>100	-
C^{001}	17.82	10.78-31.50
C^{110}	6.21	2.75-22.69
C^{101}	8.26	2.76-19.62
C^{011}	>100	-
C^{111}	0.26	0.09-0.94

Table S5.2. Total amount of protein in steady state relative to WT for WT and CAA mutant in different conditions. The table shows the total protein in steady state in the model relative to the abundance of DHHC6 observed in steady state in WT conditions. Simulations are performed for WT and CAA mutant under control conditions, after overexpression of DHHC16 and after silencing of APT2.

Protein	Condition	Amount (relative to WT control)	Range [2nd 3rd] quartiles	lower error	upper error
WT	Control	100%	[70-112]	30	12
WT	OE DHHC16	84%	[27-104]	47	20
WT	siRNA APT2	28%	[14-34.54]	14	6.52
CAA	Control	73%	[43-91]	30	18
CAA	OE DHHC16	31%	[19.60-41.50]	12	10.5
CAA	siRNA APT2	31%	[20-38]	11	7

Table S5.3 Results of stochastic simulation when APT2 is silenced. The table shows the average number of passages of a DHHC6 molecule in the different palmitoylation states when APT2 is silenced. The time spent in each state is also reported.

siRNA APT			
Species	Aver. Numb. passages /molecule	Median of the time spent per passage (h)	Median cumulative time (h)
C ⁰⁰⁰	1	0.9	0.9
C ¹⁰⁰	1	1.2	1.2
C ⁰¹⁰	<0.1	0.7	0.7
C ⁰⁰¹	<0.1	<0.1	<0.1
C ¹¹⁰	0.5	2.7	2.7
C ¹⁰¹	0.2	0.2	0.2
C ⁰¹¹	<0.1	9.9	9.9
C ¹¹¹	0.5	0.3	0.3

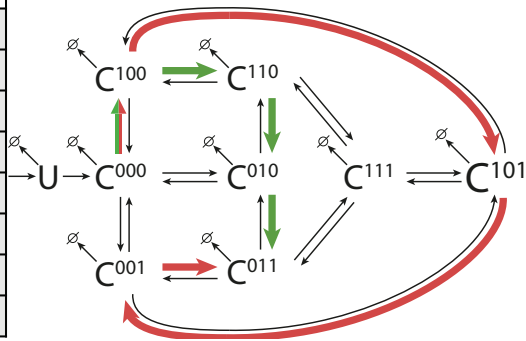


Table S5.4. Model reactions. The model of DHHC6 palmitoylation contains 35 different reactions, describing synthesis, folding, degradation, and the enzymatic reactions of DHHC6 palmitoylation/depalmitoylation. In the following table we describe in detail how the rates for those reactions are calculated.

Reaction	Forward rate	Reverse rate
7. Synthesis and folding		
$\rightarrow U$	$V_I = V_S$	-

$U \rightarrow C^{000}$	$v_2 = kf \cdot U$	-
8. Palmitoylation		
$C^{000} \rightleftharpoons C^{100}$	$v_3 = \frac{V_{f1} \cdot C^{000}}{Km_{f1} f_1}$	$v_4 = \frac{V_{b1} \cdot C^{100}}{Km_{b1} b_1}$
$C^{000} \rightleftharpoons C^{001}$	$v_5 = \frac{V_{f3} \cdot C^{000}}{Km_{f3} f_3}$	$v_6 = \frac{V_{b3} \cdot C^{001}}{Km_{b3} b_3}$
$C^{000} \rightleftharpoons C^{010}$	$v_7 = \frac{V_{f2} \cdot C^{000}}{Km_{f2} f_2}$	$v_8 = \frac{V_{b2} \cdot C^{010}}{Km_{b2} b_2}$
$C^{100} \rightleftharpoons C^{110}$	$v_9 = \frac{V_{f12} \cdot C^{100}}{Km_{f2} f_2}$	$v_{10} = \frac{V_{b12} \cdot C^{110}}{Km_{b2} b_2}$
$C^{100} \rightleftharpoons C^{101}$	$v_{11} = \frac{V_{f13} \cdot C^{100}}{Km_{f3} f_3}$	$v_{12} = \frac{V_{b13} \cdot C^{101}}{Km_{b3} b_3}$
$C^{001} \rightleftharpoons C^{101}$	$v_{13} = \frac{V_{f13} \cdot C^{001}}{Km_{f1} f_1}$	$v_{14} = \frac{V_{b13} \cdot C^{101}}{Km_{b1} b_1}$
$C^{001} \rightleftharpoons C^{011}$	$v_{15} = \frac{V_{f23} \cdot C^{001}}{Km_{f2} f_2}$	$v_{16} = \frac{V_{b23} \cdot C^{011}}{Km_{b2} b_2}$

$C^{010} \rightleftharpoons C^{110}$	$V_{17} = \frac{V_{f_{12}} \cdot C^{010}}{\frac{Km_{f_1}}{f_1}}$	$V_{18} = \frac{V_{b_{12}} \cdot C^{110}}{\frac{Km_{b_1}}{b_1}}$
$C^{010} \rightleftharpoons C^{011}$	$V_{19} = \frac{V_{f_{23}} \cdot C^{010}}{\frac{Km_{f_3}}{f_3}}$	$V_{20} = \frac{V_{b_{23}} \cdot C^{011}}{\frac{Km_{b_3}}{b_3}}$
$C^{110} \rightleftharpoons C^{111}$	$V_{21} = \frac{V_{f_{123}} \cdot C^{110}}{\frac{Km_{f_3}}{f_3}}$	$V_{22} = \frac{V_{b_{123}} \cdot C^{111}}{\frac{Km_{b_3}}{b_3}}$
$C^{101} \rightleftharpoons C^{111}$	$V_{23} = \frac{V_{f_{123}} \cdot C^{101}}{\frac{Km_{f_2}}{f_2}}$	$V_{24} = \frac{V_{b_{123}} \cdot C^{111}}{\frac{Km_{b_2}}{b_2}}$
$C^{011} \rightleftharpoons C^{111}$	$V_{25} = \frac{V_{f_{123}} \cdot C^{011}}{\frac{Km_{f_1}}{f_1}}$	$V_{26} = \frac{V_{b_{123}} \cdot C^{111}}{\frac{Km_{b_1}}{b_1}}$
9. Degradation		
$U \rightarrow \emptyset$	$V_{27} = kd_m \cdot U$	
$C^{000} \rightarrow \emptyset$	$V_{28} = kd_{c^{000}} \cdot C^{000}$	
$C^{100} \rightarrow \emptyset$	$V_{29} = kd_{c^{100}} \cdot C^{100}$	
$C^{010} \rightarrow \emptyset$	$V_{30} = kd_{c^{010}} \cdot C^{010}$	
$C^{001} \rightarrow \emptyset$	$V_{31} = kd_{c^{001}} \cdot C^{001}$	

$C^{110} \rightarrow \emptyset$	$v_{32} = kd_{c^{110}} \cdot C^{110}$	
$C^{101} \rightarrow \emptyset$	$v_{33} = kd_{c^{101}} \cdot C^{101}$	
$C^{011} \rightarrow \emptyset$	$v_{34} = kd_{c^{011}} \cdot C^{011}$	
$C^{111} \rightarrow \emptyset$	$v_{35} = kd_{c^{111}} \cdot C^{111}$	

Where the competition terms f and b in SI Table are defined as:

f_1	$1 + \frac{C^{000}}{Km_{f_1}} + \frac{C^{000}}{Km_{f_2}} + \frac{C^{100}}{Km_{f_3}} + \frac{C^{100}}{Km_{f_1}} + \frac{C^{010}}{Km_{f_1}} + \frac{C^{010}}{Km_{f_3}} + \frac{C^{001}}{Km_{f_1}} + \frac{C^{001}}{Km_{f_2}} + \frac{C^{110}}{Km_{f_3}} + \frac{C^{101}}{Km_{f_2}} + \frac{C^{011}}{Km_{f_1}} + \frac{DHHHC_{16L}}{Km_{f_1}}$
f_2	$1 + \frac{C^{000}}{Km_{f_1}} + \frac{C^{000}}{Km_{f_2}} + \frac{C^{100}}{Km_{f_3}} + \frac{C^{100}}{Km_{f_1}} + \frac{C^{010}}{Km_{f_1}} + \frac{C^{010}}{Km_{f_3}} + \frac{C^{001}}{Km_{f_1}} + \frac{C^{001}}{Km_{f_2}} + \frac{C^{110}}{Km_{f_3}} + \frac{C^{101}}{Km_{f_2}} + \frac{C^{011}}{Km_{f_1}} + \frac{DHHHC_{16L}}{Km_{f_2}}$
f_3	$1 + \frac{C^{000}}{Km_{f_1}} + \frac{C^{000}}{Km_{f_2}} + \frac{C^{100}}{Km_{f_3}} + \frac{C^{100}}{Km_{f_1}} + \frac{C^{010}}{Km_{f_1}} + \frac{C^{010}}{Km_{f_3}} + \frac{C^{001}}{Km_{f_1}} + \frac{C^{001}}{Km_{f_2}} + \frac{C^{110}}{Km_{f_3}} + \frac{C^{101}}{Km_{f_2}} + \frac{C^{011}}{Km_{f_1}} + \frac{DHHHC_{16L}}{Km_{f_3}}$
b_1	$1 + \frac{C^{111}}{Km_{b_3}} + \frac{C^{111}}{Km_{b_2}} + \frac{C^{111}}{Km_{b_1}} + \frac{C^{100}}{Km_{b_1}} + \frac{C^{010}}{Km_{b_2}} + \frac{C^{001}}{Km_{b_3}} + \frac{C^{110}}{Km_{b_1}} + \frac{C^{101}}{Km_{b_1}} + \frac{C^{011}}{Km_{b_2}} + \frac{APT_{2L}}{Km_{b_1}}$
b_2	$1 + \frac{C^{111}}{Km_{b_3}} + \frac{C^{111}}{Km_{b_2}} + \frac{C^{111}}{Km_{b_1}} + \frac{C^{100}}{Km_{b_1}} + \frac{C^{010}}{Km_{b_2}} + \frac{C^{001}}{Km_{b_3}} + \frac{C^{110}}{Km_{b_1}} + \frac{C^{101}}{Km_{b_1}} + \frac{C^{011}}{Km_{b_2}} + \frac{APT_{2L}}{Km_{b_2}}$
b_3	$1 + \frac{C^{111}}{Km_{b_3}} + \frac{C^{111}}{Km_{b_2}} + \frac{C^{111}}{Km_{b_1}} + \frac{C^{100}}{Km_{b_1}} + \frac{C^{010}}{Km_{b_2}} + \frac{C^{001}}{Km_{b_3}} + \frac{C^{110}}{Km_{b_1}} + \frac{C^{101}}{Km_{b_1}} + \frac{C^{011}}{Km_{b_2}} + \frac{APT_{2L}}{Km_{b_3}}$

Table S5.5. Model parameters. The output of GA is a set of optimal solutions, where a solution is a complete set of parameter needed to perform model simulations. From this set we extracted a sub-set of 152 solutions that obtained a GA score better than a set threshold for each objective. During the analysis the model was simulated for each set of parameters of the sub-set. We then reported in this paper the mean of the outputs along with the 1st and 3rd quartile of their distribution.

Parameter	Median Value	Range
$v_s [C]^a/[T]^b$	$3.00 \cdot 10^{-3}$	$[2.59 \cdot 10^{-3} - 3.31 \cdot 10^{-3}]$
$k_f 1/[T]$	$4.20 \cdot 10^{-1}$	$[2.98 \cdot 10^{-1} - 6.69 \cdot 10^{-1}]$
$V_{f1} [C]/[T]$	6.03	$[4.54 - 6.92]$
$Km_{f1} [C]$	$2.02 \cdot 10^{-2}$	$[1.62 \cdot 10^{-2} - 2.48 \cdot 10^{-2}]$
$V_{f2} [C]/[T]$	$4.21 \cdot 10^{-4}$	$[1.18 \cdot 10^{-4} - 9.89 \cdot 10^{-4}]$
$Km_{f2} [C]$	$1.33 \cdot 10^{-5}$	$[9.40 \cdot 10^{-6} - 2.16 \cdot 10^{-5}]$
$V_{f3} [C]/[T]$	$8.07 \cdot 10^{-2}$	$[4.30 \cdot 10^{-2} - 1.49 \cdot 10^{-1}]$
$Km_{f3} [C]$	$1.06 \cdot 10^{-2}$	$[8.21 \cdot 10^{-3} - 1.22 \cdot 10^{-2}]$
$V_{b3} [C]/[T]$	$8.12 \cdot 10^{-3}$	$[5.07 \cdot 10^{-3} - 1.13 \cdot 10^{-2}]$
$Km_{b3} [C]$	$1.63 \cdot 10^{-3}$	$[7.99 \cdot 10^{-4} - 3.22 \cdot 10^{-3}]$
$V_{b2} [C]/[T]$	$5.86 \cdot 10^{-2}$	$[4.52 \cdot 10^{-2} - 7.44 \cdot 10^{-2}]$
$Km_{b2} [C]$	$1.74 \cdot 10^{-2}$	$[1.18 \cdot 10^{-2} - 2.29 \cdot 10^{-2}]$
$V_{b1} [C]/[T]$	$1.18 \cdot 10^{-1}$	$[9.04 \cdot 10^{-2} - 1.42 \cdot 10^{-1}]$
$Km_{b1} [C]$	$3.71 \cdot 10^{-3}$	$[2.27 \cdot 10^{-3} - 6.13 \cdot 10^{-3}]$
$kd_m 1/[T]$	$2.56 \cdot 10^{-1}$	$[2.16 \cdot 10^{-1} - 3.50 \cdot 10^{-1}]$
$kd_{c^{000}} 1/[T]$	$1.80 \cdot 10^{-2}$	$[1.47 \cdot 10^{-2} - 2.72 \cdot 10^{-2}]$
$kd_{c^{100}} 1/[T]$	$1.33 \cdot 10^{-1}$	$[8.81 \cdot 10^{-2} - 1.97 \cdot 10^{-1}]$
$kd_{c^{010}} 1/[T]$	$4.82 \cdot 10^{-3}$	$[1.58 \cdot 10^{-3} - 1.32 \cdot 10^{-2}]$
$kd_{c^{001}} 1/[T]$	$3.89 \cdot 10^{-2}$	$[2.07 \cdot 10^{-2} - 6.04 \cdot 10^{-2}]$
$kd_{c^{110}} 1/[T]$	$1.28 \cdot 10^{-1}$	$[3.66 \cdot 10^{-2} - 2.50 \cdot 10^{-1}]$
$kd_{c^{101}} 1/[T]$	$5.59 \cdot 10^{-2}$	$[2.74 \cdot 10^{-2} - 2.41 \cdot 10^{-1}]$
$kd_{c^{011}} 1/[T]$	$3.42 \cdot 10^{-4}$	$[6.46 \cdot 10^{-5} - 1.36 \cdot 10^{-3}]$
$kd_{c^{111}} 1/[T]$	2.32	$[7.14 \cdot 10^{-1} - 7.31]$
$V_{f12} [C]/[T]$	$6.43 \cdot 10^{-2}$	$[3.36 \cdot 10^{-2} - 1.07 \cdot 10^{-1}]$
$V_{f13} [C]/[T]$	$6.04 \cdot 10^{-1}$	$[1.54 \cdot 10^{-1} - 1.36]$
$V_{f23} [C]/[T]$	2.89	$[1.58 - 4.74]$
$V_{f123} [C]/[T]$	$5.61 \cdot 10^{-1}$	$[1.37 \cdot 10^{-1} - 2.43]$

V_{b12} [C]/[T]	1.83	$[7.00 \cdot 10^{-1} - 5.64]$
V_{b13} [C]/[T]	1.40	$[5.01 \cdot 10^{-1} - 5.28]$
V_{b23} [C]/[T]	$9.49 \cdot 10^{-2}$	$[5.60 \cdot 10^{-2} - 1.54 \cdot 10^{-1}]$
V_{b123} [C]/[T]	$9.35 \cdot 10^{-1}$	$[2.09 \cdot 10^{-1} - 4.19]$
$DHHC_{16t}$ [C]	$1.73 \cdot 10^{-1}$	$[1.05 \cdot 10^{-1} - 3.10 \cdot 10^{-1}]$
APT_{2t} [C]	$1.85 \cdot 10^{-1}$	$[1.42 \cdot 10^{-1} - 2.20 \cdot 10^{-1}]$

Table S5.6. Mass balance equations. The following table describes the mass balance for each of the species of DHHC6 model. The rates of the mass balance of each state are described in detail in S1 Table.

Mass balance	Mass balance
$\frac{dU}{dt} = v1 - v2 - v27$	$\frac{dC^{110}}{dt} = v9 + v17 + v22 - v10 - v18 - v21 - v32$
$\frac{dC^{000}}{dt} = v2 + v4 + v6 + v8 - v3 - v5 - v7 - v28$	$\frac{dC^{101}}{dt} = v11 + v13 + v24 - v12 - v14 - v23 - v33$
$\frac{dC^{100}}{dt} = v3 + v10 + v12 - v4 - v9 - v11 - v29$	$\frac{dC^{011}}{dt} = v15 + v19 + v26 - v16 - v20 - v25 - v34$
$\frac{dC^{010}}{dt} = v7 + v18 + v20 - v8 - v17 - v19 - v30$	$\frac{dC^{111}}{dt} = v21 + v23 + v25 - v22 - v26 - v24 - v35$
$\frac{dC^{001}}{dt} = v5 + v14 + v16 - v6 - v13 - v15 - v31$	

Table S5.7. Parameters used for stochastic simulations. The following parameters were obtained through the conversion of the deterministic parameters estimated by the GA (see [39]).

Parameter	Value	Units
v_s	$8.76 \cdot 10^1$	molecules/h
k_f	$4.20 \cdot 10^{-1}$	1/h
V_{f1}	$1.76 \cdot 10^5$	molecules/h
Km_{f1}	$5.91 \cdot 10^2$	Molecules
V_{f2}	$1.23 \cdot 10^1$	molecules/h
Km_{f2}	$3.89 \cdot 10^{-1}$	Molecules
V_{f3}	$2.36 \cdot 10^3$	molecules/h
Km_{f3}	$3.09 \cdot 10^2$	molecules
V_{b3}	$1.58 \cdot 10^2$	molecules/h
Km_{b3}	$4.77 \cdot 10^1$	Molecules
V_{b2}	$1.71 \cdot 10^3$	molecules/h
Km_{b2}	$5.07 \cdot 10^2$	Molecules
V_{b1}	$3.46 \cdot 10^3$	molecules/h
Km_{b1}	$1.08 \cdot 10^2$	Molecules
kd_m	$2.56 \cdot 10^{-1}$	1/h
$kd_{c^{000}}$	$1.80 \cdot 10^{-2}$	1/h
$kd_{c^{100}}$	$1.33 \cdot 10^{-1}$	1/h
$kd_{c^{010}}$	$4.82 \cdot 10^{-3}$	1/h
$kd_{c^{001}}$	$3.89 \cdot 10^{-2}$	1/h
$kd_{c^{110}}$	$1.28 \cdot 10^{-1}$	1/h
$kd_{c^{101}}$	$5.59 \cdot 10^{-2}$	1/h
$kd_{c^{011}}$	$3.42 \cdot 10^{-3}$	1/h
$kd_{c^{111}}$	2.32	1/h
V_{f12}	$1.88 \cdot 10^3$	molecules/h
V_{f13}	$1.76 \cdot 10^4$	molecules/h
V_{f23}	$8.45 \cdot 10^4$	molecules/h
V_{f123}	$1.64 \cdot 10^4$	molecules/h
V_{b12}	$5.35 \cdot 10^4$	molecules/h

$V_{b_{13}}$	$4.09 \cdot 10^4$	molecules/h
$V_{b_{23}}$	$2.77 \cdot 10^3$	molecules/h
$V_{b_{123}}$	$2.73 \cdot 10^4$	molecules/h
$DHHC_{16t}$	$5.06 \cdot 10^3$	molecules
APT_{2t}	$5.39 \cdot 10^3$	Molecules

Table S5.8 Stoichiometric matrix used for stochastic simulations. In the following table we define the stoichiometry and the directionality of the reactions of the model. Each reaction has a directionality that define which are the reagents and which are the products. In this table each line represents a model reaction, while in the columns we find all the states of the model. For each reaction the states of the model that take part to the reaction as reagents are marked with -1, while the states that participate as products are marked with 1. The matrix that is formed in this way allow to attribute the correct directionality to model reactions during the calculation of the mass balance for each state of the model.

	U	C000	C100	C010	C001	C110	C101	C011	C111	U1	C0001	C1001	C0011	C1101	C1011	C0111	C1111	C0001_#0	C1001_#0	C0011_#0	C1101_#0	C1011_#0	C0111_#0	C1111_#0
U -> U	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C000 -> C100	0	-1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C000 -> C010	0	-1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C000 -> C001	0	-1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C100 -> C110	0	0	-1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C100 -> C101	0	0	-1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C100 -> C010	0	0	-1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C010 -> C110	0	0	0	-1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C010 -> C011	0	0	0	-1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C010 -> C100	0	0	0	-1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C001 -> C011	0	0	0	0	-1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C001 -> C000	0	0	0	0	-1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C001 -> C100	0	0	0	0	-1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C100 -> C111	0	0	1	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C100 -> C011	0	0	1	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C100 -> C001	0	0	1	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C010 -> C111	0	0	0	1	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C010 -> C101	0	0	0	1	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C010 -> C011	0	0	0	1	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C110 -> C111	0	0	0	0	1	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C110 -> C100	0	0	0	0	1	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C110 -> C010	0	0	0	0	1	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0
C111 -> C110	0	0	0	0	0	1	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0
C111 -> C100	0	0	0	0	0	1	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0
C111 -> C010	0	0	0	0	0	1	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0
U1 -> U1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C0001 -> C1001	0	0	0	0	0	0	0	0	0	0	-1	1	0	0	0	0	0	0	0	0	0	0	0	0
C0001 -> C0101	0	0	0	0	0	0	0	0	0	0	-1	0	1	0	0	0	0	0	0	0	0	0	0	0
C0001 -> C0011	0	0	0	0	0	0	0	0	0	0	0	-1	0	1	0	0	0	0	0	0	0	0	0	0
C1001 -> C1101	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1001 -> C1011	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1001 -> C0101	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C0101 -> C1101	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C0101 -> C1011	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C0101 -> C0111	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1101 -> C1111	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1101 -> C1100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1101 -> C1000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1101 -> C0100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1101 -> C1110	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1101 -> C1010	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1101 -> C0110	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1111 -> C1110	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1111 -> C1100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1111 -> C1000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1111 -> C0100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1111 -> C1110	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1111 -> C1100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1111 -> C1000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1111 -> C0100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1111 -> C1110	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1111 -> C1100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1111 -> C1000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1111 -> C0100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1111 -> C1110	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1111 -> C1100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1111 -> C1000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1111 -> C0100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1111 -> C1110	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1111 -> C1100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1111 -> C1000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1111 -> C0100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1111 -> C1110	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1111 -> C1100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1111 -> C1000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1111 -> C0100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1111 -> C1110	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1111 -> C1100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1111 -> C1000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1111 -> C0100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1111 -> C1110	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1111 -> C1100	0	0	0																					

Table 5.9. Propensity function used for stochastic simulations. In the first column of the table, each line describes a reaction of the model. To each reaction is associated a rate, in the second column, that describes the probability of that reaction to happen at each time step of the stochastic simulation.

Reaction	Propensity
$\rightarrow U$	vs
$U \rightarrow C^{000}$	$kf \cdot U$
$C^{000} \rightarrow C^{100}$	$\frac{V_{f1} \cdot C^{000}}{\frac{Km_{f1}}{f_1}}$
$C^{000} \rightarrow C^{010}$	$\frac{V_{f2} \cdot C^{000}}{\frac{Km_{f2}}{f_2}}$
$C^{000} \rightarrow C^{001}$	$\frac{V_{f3} \cdot C^{000}}{\frac{Km_{f3}}{f_3}}$
$C^{000} \rightarrow \emptyset$	$kd_{c^{000}} \cdot C^{000}$
$C^{100} \rightarrow C^{110}$	$\frac{V_{f12} \cdot C^{100}}{\frac{Km_{f2}}{f_2}}$
$C^{100} \rightarrow C^{101}$	$\frac{V_{f13} \cdot C^{100}}{\frac{Km_{f3}}{f_3}}$

$C^{100} \rightarrow C^{000}$	$\frac{V_{b1} \cdot C^{100}}{\frac{Km_{b1}}{b_1}}$
$C^{100} \rightarrow \emptyset$	$kd_{c^{100}} \cdot C^{100}$
$C^{010} \rightarrow C^{110}$	$\frac{V_{f12} \cdot C^{010}}{\frac{Km_{f1}}{f_1}}$
$C^{010} \rightarrow C^{011}$	$\frac{V_{f23} \cdot C^{010}}{\frac{Km_{f3}}{f_3}}$
$C^{010} \rightarrow C^{000}$	$\frac{V_{b2} \cdot C^{010}}{\frac{Km_{b2}}{b_2}}$
$C^{010} \rightarrow \emptyset$	$kd_{c^{010}} \cdot C^{010}$
$C^{001} \rightarrow C^{101}$	$\frac{V_{f13} \cdot C^{001}}{\frac{Km_{f1}}{f_1}}$
$C^{001} \rightarrow C^{011}$	$\frac{V_{f23} \cdot C^{001}}{\frac{Km_{f2}}{f_2}}$
$C^{001} \rightarrow C^{000}$	$\frac{V_{b3} \cdot C^{001}}{\frac{Km_{b3}}{b_3}}$
$C^{001} \rightarrow \emptyset$	$kd_{c^{001}} \cdot C^{001}$

$C^{110} \rightarrow C^{111}$	$\frac{V_{f_{123}} \cdot C^{110}}{\frac{Km_{f_3}}{f_3}}$
$C^{110} \rightarrow C^{100}$	$\frac{V_{b_{12}} \cdot C^{110}}{\frac{Km_{b_2}}{b_2}}$
$C^{110} \rightarrow C^{010}$	$\frac{V_{b_{12}} \cdot C^{110}}{\frac{Km_{b_1}}{b_1}}$
$C^{110} \rightarrow \emptyset$	$kd_{c^{110}} \cdot C^{110}$
$C^{101} \rightarrow C^{111}$	$\frac{V_{f_{123}} \cdot C^{101}}{\frac{Km_{f_2}}{f_2}}$
$C^{101} \rightarrow C^{100}$	$\frac{V_{b_{13}} \cdot C^{101}}{\frac{Km_{b_3}}{b_3}}$
$C^{101} \rightarrow C^{001}$	$\frac{V_{b_{13}} \cdot C^{101}}{\frac{Km_{b_1}}{b_1}}$
$C^{101} \rightarrow \emptyset$	$kd_{c^{101}} \cdot C^{101}$
$C^{011} \rightarrow C^{111}$	$\frac{V_{f_{123}} \cdot C^{011}}{\frac{Km_{f_1}}{f_1}}$

$C^{011} \rightarrow C^{010}$	$\frac{V_{b_{23}} \cdot C^{011}}{\frac{Km_{b_3}}{b_3}}$
$C^{011} \rightarrow C^{001}$	$\frac{V_{b_{23}} \cdot C^{011}}{\frac{Km_{b_2}}{b_2}}$
$C^{011} \rightarrow \emptyset$	$kd_{c^{011}} \cdot C^{011}$
$C^{111} \rightarrow C^{110}$	$\frac{V_{b_{123}} \cdot C^{111}}{\frac{Km_{b_3}}{b_3}}$
$C^{111} \rightarrow C^{101}$	$\frac{V_{b_{123}} \cdot C^{111}}{\frac{Km_{b_2}}{b_2}}$
$C^{111} \rightarrow C^{011}$	$\frac{V_{b_{123}} \cdot C^{111}}{\frac{Km_{b_1}}{b_1}}$
$C^{111} \rightarrow \emptyset$	$kd_{c^{111}} \cdot C^{111}$
$\rightarrow U^*$	vs
$U^* \rightarrow C^{000^*}$	$kf \cdot U^*$
$C^{000^*} \rightarrow C^{100^*}$	$\frac{V_{f_1} \cdot C^{000^*}}{\frac{Km_{f_1}}{f_1}}$

$C^{000*} \rightarrow C^{010*}$	$\frac{V_{f_2} \cdot C^{000*}}{\frac{Km_{f_2}}{f_2}}$
$C^{000*} \rightarrow C^{001*}$	$\frac{V_{f_3} \cdot C^{000*}}{\frac{Km_{f_3}}{f_3}}$
$C^{000*} \rightarrow \emptyset$	$kd_{c^{000}} \cdot C^{000*}$
$C^{100*} \rightarrow C^{110*}$	$\frac{V_{f_{12}} \cdot C^{100*}}{\frac{Km_{f_2}}{f_2}}$
$C^{100*} \rightarrow C^{101*}$	$\frac{V_{f_{13}} \cdot C^{100*}}{\frac{Km_{f_3}}{f_3}}$
$C^{100*} \rightarrow C^{000*}$	$\frac{V_{b_1} \cdot C^{100*}}{\frac{Km_{b_1}}{b_1}}$
$C^{100*} \rightarrow \emptyset$	$kd_{c^{100}} \cdot C^{100*}$
$C^{010*} \rightarrow C^{110*}$	$\frac{V_{f_{12}} \cdot C^{010*}}{\frac{Km_{f_1}}{f_1}}$
$C^{010*} \rightarrow C^{011*}$	$\frac{V_{f_{23}} \cdot C^{010*}}{\frac{Km_{f_3}}{f_3}}$

$C^{010*} \rightarrow C^{000*}$	$\frac{V_{b_2} \cdot C^{010*}}{\frac{Km_{b_2}}{b_2}}$
$C^{010*} \rightarrow \emptyset$	$kd_{c^{010}} \cdot C^{010*}$
$C^{001*} \rightarrow C^{101*}$	$\frac{V_{f_{13}} \cdot C^{001*}}{\frac{Km_{f_1}}{f_1}}$
$C^{001*} \rightarrow C^{011*}$	$\frac{V_{f_{23}} \cdot C^{001*}}{\frac{Km_{f_2}}{f_2}}$
$C^{001*} \rightarrow C^{000*}$	$\frac{V_{b_3} \cdot C^{001*}}{\frac{Km_{b_3}}{b_3}}$
$C^{001*} \rightarrow \emptyset$	$kd_{c^{001}} \cdot C^{001*}$
$C^{110*} \rightarrow C^{111*}$	$\frac{V_{f_{123}} \cdot C^{110*}}{\frac{Km_{f_3}}{f_3}}$
$C^{110*} \rightarrow C^{100*}$	$\frac{V_{b_{12}} \cdot C^{110*}}{\frac{Km_{b_2}}{b_2}}$
$C^{110*} \rightarrow C^{010*}$	$\frac{V_{b_{12}} \cdot C^{110*}}{\frac{Km_{b_1}}{b_1}}$
$C^{110*} \rightarrow \emptyset$	$kd_{c^{110}} \cdot C^{110*}$

$C^{101*} \rightarrow C^{111*}$	$\frac{V_{f_{123}} \cdot C^{101*}}{\frac{Km_{f_2}}{f_2}}$
$C^{101*} \rightarrow C^{100*}$	$\frac{V_{b_{13}} \cdot C^{101*}}{\frac{Km_{b_3}}{b_3}}$
$C^{101*} \rightarrow C^{001*}$	$\frac{V_{b_{13}} \cdot C^{101*}}{\frac{Km_{b_1}}{b_1}}$
$C^{101*} \rightarrow \emptyset$	$kd_{c^{101}} \cdot C^{101*}$
$C^{011*} \rightarrow C^{111*}$	$\frac{V_{f_{123}} \cdot C^{011*}}{\frac{Km_{f_1}}{f_1}}$
$C^{011*} \rightarrow C^{010*}$	$\frac{V_{b_{23}} \cdot C^{011*}}{\frac{Km_{b_2}}{b_2}}$
$C^{011*} \rightarrow C^{001*}$	$\frac{V_{b_{23}} \cdot C^{011*}}{\frac{Km_{b_3}}{b_3}}$
$C^{011*} \rightarrow \emptyset$	$kd_{c^{011}} \cdot C^{011*}$
$C^{111*} \rightarrow C^{110*}$	$\frac{V_{b_{123}} \cdot C^{111*}}{\frac{Km_{b_3}}{b_3}}$

$C^{111*} \rightarrow C^{101*}$	$\frac{V_{b_{123}} \cdot C^{111*}}{\frac{Km_{b_2}}{b_2}}$
$C^{111*} \rightarrow C^{011*}$	$\frac{V_{b_{123}} \cdot C^{111*}}{\frac{Km_{b_1}}{b_1}}$
$C^{111*} \rightarrow \emptyset$	$kd_{c^{111}} \cdot C^{111*}$

Where the competition terms f and b in SI Table are defined as:

f_1	$1 + \frac{C^{000}}{Km_{f_1}} + \frac{C^{100}}{Km_{f_2}} + \frac{C^{010}}{Km_{f_3}} + \frac{C^{100}}{Km_{f_1}} + \frac{C^{010}}{Km_{f_3}} + \frac{C^{001}}{Km_{f_1}} + \frac{C^{101}}{Km_{f_2}} + \frac{C^{011}}{Km_{f_1}}$ $+ \frac{C^{000}}{Km_{f_1}} + \frac{C^{100}}{Km_{f_2}} + \frac{C^{010}}{Km_{f_3}} + \frac{C^{100}}{Km_{f_1}} + \frac{C^{010}}{Km_{f_3}} + \frac{C^{001}}{Km_{f_1}} + \frac{C^{101}}{Km_{f_2}} + \frac{C^{011}}{Km_{f_1}} + \frac{DHHc_{16f}}{Km_{f_1}}$
f_2	$1 + \frac{C^{000}}{Km_{f_1}} + \frac{C^{100}}{Km_{f_2}} + \frac{C^{010}}{Km_{f_3}} + \frac{C^{100}}{Km_{f_1}} + \frac{C^{010}}{Km_{f_3}} + \frac{C^{001}}{Km_{f_1}} + \frac{C^{101}}{Km_{f_2}} + \frac{C^{011}}{Km_{f_1}}$ $+ \frac{C^{000}}{Km_{f_1}} + \frac{C^{100}}{Km_{f_2}} + \frac{C^{010}}{Km_{f_3}} + \frac{C^{100}}{Km_{f_1}} + \frac{C^{010}}{Km_{f_3}} + \frac{C^{001}}{Km_{f_1}} + \frac{C^{101}}{Km_{f_2}} + \frac{C^{011}}{Km_{f_1}} + \frac{DHHc_{16f}}{Km_{f_2}}$
f_3	$1 + \frac{C^{000}}{Km_{f_1}} + \frac{C^{100}}{Km_{f_2}} + \frac{C^{010}}{Km_{f_3}} + \frac{C^{100}}{Km_{f_1}} + \frac{C^{010}}{Km_{f_3}} + \frac{C^{001}}{Km_{f_1}} + \frac{C^{101}}{Km_{f_2}} + \frac{C^{011}}{Km_{f_1}}$ $+ \frac{C^{000}}{Km_{f_1}} + \frac{C^{100}}{Km_{f_2}} + \frac{C^{010}}{Km_{f_3}} + \frac{C^{100}}{Km_{f_1}} + \frac{C^{010}}{Km_{f_3}} + \frac{C^{001}}{Km_{f_1}} + \frac{C^{101}}{Km_{f_2}} + \frac{C^{011}}{Km_{f_1}} + \frac{DHHc_{16f}}{Km_{f_3}}$
b_1	$1 + \frac{C^{111}}{Km_{b_3}} + \frac{C^{111}}{Km_{b_2}} + \frac{C^{111}}{Km_{b_1}} + \frac{C^{100}}{Km_{b_1}} + \frac{C^{010}}{Km_{b_2}} + \frac{C^{001}}{Km_{b_3}} + \frac{C^{110}}{Km_{b_1}} + \frac{C^{101}}{Km_{b_2}} + \frac{C^{011}}{Km_{b_1}}$ $+ \frac{C^{111*}}{Km_{b_1}} + \frac{C^{100*}}{Km_{b_2}} + \frac{C^{010*}}{Km_{b_3}} + \frac{C^{110*}}{Km_{b_1}} + \frac{C^{101*}}{Km_{b_2}} + \frac{C^{011*}}{Km_{b_3}} + \frac{APT}{2t} + \frac{Km_{b_2}}{Km_{b_1}}$

b_2	$1 + \frac{C^{111}}{Km_{b3}} + \frac{C^{111}}{Km_{b2}} + \frac{C^{111}}{Km_{b1}} + \frac{C^{100}}{Km_{b1}} + \frac{C^{010}}{Km_{b2}} + \frac{C^{001}}{Km_{b3}} + \frac{C^{110}}{Km_{b1}} + \frac{C^{110}}{Km_{b2}} + \frac{C^{101}}{Km_{b1}} + \frac{C^{101}}{Km_{b3}} + \frac{C^{011}}{Km_{b3}} + \frac{C^{011}}{Km_{b2}}$ $+ \frac{C^{111*}}{Km_{b1}} + \frac{C^{100*}}{Km_{b1}} + \frac{C^{010*}}{Km_{b2}} + \frac{C^{001*}}{Km_{b3}} + \frac{C^{110*}}{Km_{b1}} + \frac{C^{110*}}{Km_{b2}} + \frac{C^{101*}}{Km_{b1}} + \frac{C^{101*}}{Km_{b3}} + \frac{C^{011*}}{Km_{b3}} + \frac{APT}{2t} + \frac{APT}{2t} + \frac{Km_{b2}}{Km_{b2}}$
b_3	$1 + \frac{C^{111}}{Km_{b3}} + \frac{C^{111}}{Km_{b2}} + \frac{C^{111}}{Km_{b1}} + \frac{C^{100}}{Km_{b1}} + \frac{C^{010}}{Km_{b2}} + \frac{C^{001}}{Km_{b3}} + \frac{C^{110}}{Km_{b1}} + \frac{C^{110}}{Km_{b2}} + \frac{C^{101}}{Km_{b1}} + \frac{C^{101}}{Km_{b3}} + \frac{C^{011}}{Km_{b3}} + \frac{C^{011}}{Km_{b2}}$ $+ \frac{C^{111*}}{Km_{b1}} + \frac{C^{100*}}{Km_{b1}} + \frac{C^{010*}}{Km_{b2}} + \frac{C^{001*}}{Km_{b3}} + \frac{C^{110*}}{Km_{b1}} + \frac{C^{110*}}{Km_{b2}} + \frac{C^{101*}}{Km_{b1}} + \frac{C^{101*}}{Km_{b3}} + \frac{C^{011*}}{Km_{b3}} + \frac{APT}{2t} + \frac{APT}{2t} + \frac{Km_{b3}}{Km_{b3}}$

Chapter 6 - Reconstruction of a palmitoylation network

6.1 Introduction

In this chapter we show how we reconstructed the palmitoylation cascade of calnexin. This model was created to study the consequences of palmitoylation at the system level. In particular, we will first introduce the methods that we choose to design the network and the tools that have been developed for its analysis. Then we show some preliminary results that illustrate the capabilities of the model. The model and all the methods presented in this chapter are particularly valuable since they were developed thinking to a possible future expansion of the network, including more DHHCs and palmitoylation substrates.

6.2 Network design

The choice of modelling the calnexin palmitoylation cascade to study the network of palmitoylation seemed natural to us, due to the high availability of experimental data and the fact that we had already modelled two of the main proteins participating in this network. The network of palmitoylation, as we designed it, is composed of 5 proteins: DHHCl6, DHHC6, calnexin, APT1 and APT2 (Figure 6.1).

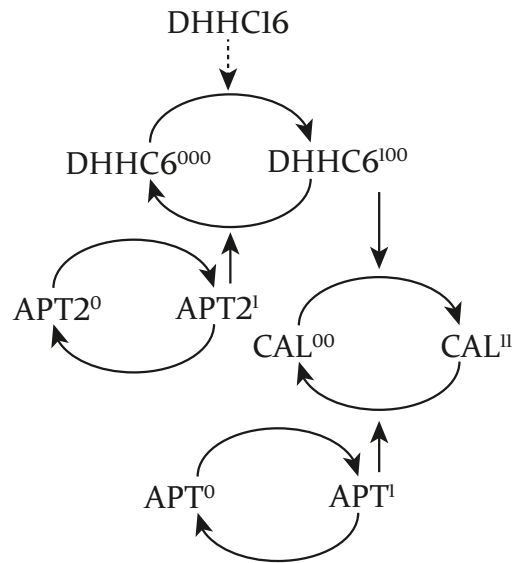


Figure 6.1. Representation of calnexin palmitoylation cascade. This network is composed of 5 proteins and describe the palmitoylation cascade of calnexin, which is palmitoylated by DHHc6 and depalmitoylated by APT1. DHHc6 is itself palmitoylated by DHHc16 and depalmitoylated by APT2. Both APT1 and APT2 are palmitoylated but it is not known who regulates their cycle of palmitoylation/depalmitoylation, therefore their modification cycle was described with simple mass-action terms.

In this model calnexin (CAL) can be palmitoylated on 2 sites by DHHc6, while the opposite reaction of depalmitoylation is catalysed by APT1. Both APT1 and DHHc6 needs to be palmitoylated to be active; while for APT1 we don't know the enzyme responsible for its modification, in the case of DHHc6 it is DHHc16. DHHc16 can palmitoylate DHHc6 on three different sites, and the opposite reaction of palmitate removal is performed by APT2. In this network we have terms that describe synthesis of calnexin and DHHc6, as well as parameters that define the degradation rates of DHHc6 and calnexin in each of their palmitoylation states. As stated above we don't have information on the palmitoylation/depalmitoylation cycles of APT1 and APT2. The enzymes involved in those reactions are still unknown, as consequence, their cycle of modification isn't modelled as for DHHc6 or calnexin (see Chapter 3 -). Instead we decided to model palmitoylation and depalmitoylation of APTs with simple mass-action kinetics. We decided to include the cycle of modification of APTs even if we didn't have all the kinetic information because the previous models have highlighted activity of APT is an important parameter in the regulation of palmitoylation cycles. Therefore, we thought that being able to tune APTs activity based on their palmitoylation levels would help to better characterize the regulation of

palmitoylation cycles. Similarly to APTs, very little information are available on DHHC16. It is not palmitoylated, and we don't know if its activity is regulated. Moreover, very little is known about regulation of its synthesis and degradation. Therefore, in our model DHHC16 is implemented with fixed concentration, and the only reaction in which this enzyme participate is catalysis of DHHC6 palmitoylation. The amount of activity of DHHC16 can still be tuned by varying its concentration.

Once we defined all the participant in the network, the interaction between them and the types of kinetics to adopt for each reaction, we had to deal with another major challenge. Even if this is a small network, its complexity is very high. The complexity of this model come from two main source, one being the number of palmitoylation sites present on each protein in the network. Multiple palmitoylation sites implies that a protein can exist in many states characterized by different palmitoylation patterns. These different protein forms can interact with other proteins which can have different states as well, exponentially increasing the number of possible interaction and complex formation based on the number of sites of each protein. The second source of complexity is due to the implementation of the *in-silico* labelling technique used for model calibration and analysis. This is because *in-silico* labelling require the definition of a new set of species (the labelled protein), which double the number of species that can exist in the model. To make an example we analysed how the number of species and interactions between calnexin and DHHC6 increase when we consider multiple sites and labelling Figure 6.2.

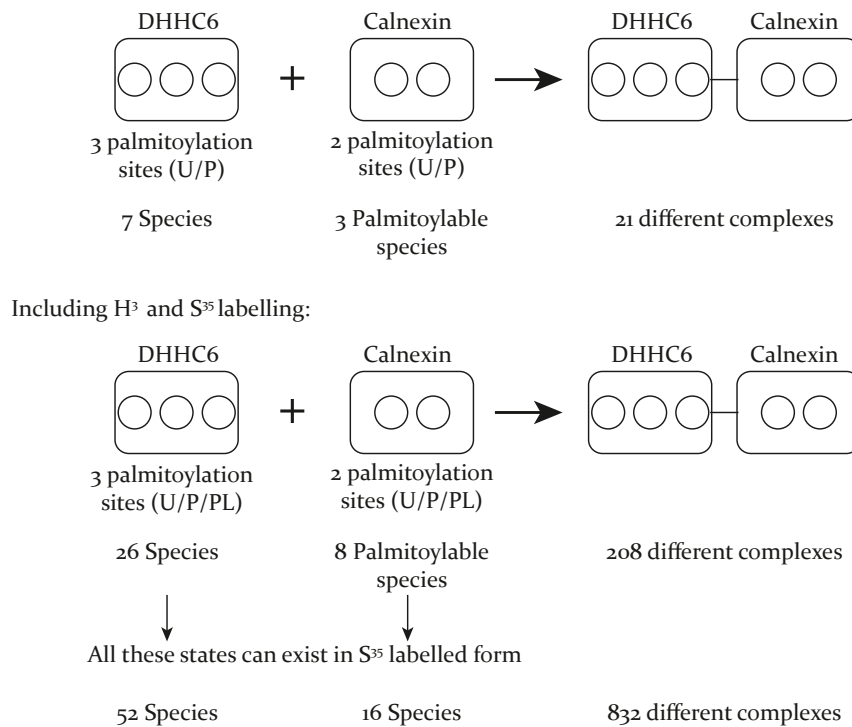


Figure 6.2. Illustration of the complexity rising from the presence of multiple modification sites on protein and the implementation of *in-silico* labelling techniques.

Figure 6.2 show that the number of species increase exponentially with the number of sites and the labelling techniques implemented. If we consider all the proteins included in the model (Figure 6.1) in all the possible states, we will have to write 705 ODEs and 2818 reaction fluxes. Writing this kind of system by hand would be extremely time consuming and it would also be difficult to ensure the correctness of the model. Therefore, we had to find an automated and more efficient method allowing us to generate the system of ODEs describing this model. For this purpose we used a particular kind of modelling approach called “Rule-based modelling” [140,141]. This kind of modelling allows to specify a set of rules that define how reactions take place, which species are involved, and which parameters define the kinetics of reaction. The rule-set can then be translated into an ODEs model automatically. This approach hugely simplifies the design of the model since the rule-set is simpler than the model that it describes. This is especially true for modification cascades, which are normally made of repeated reaction schemes, making possible to use the same set of

rules to describe different parts of the model. Moreover, the modularity of this approach made it particularly indicated in our case, since we aim to expand the network in the future. The rules that we defined to describe the palmitoylation cascade of calnexin, may be used to describe different part of the palmitoylation network, like for example the modification other DHHC6 substrates, like for example climp63, a transmembrane protein that work as receptor on the plasma membrane or the E3 ubiquitin ligase gp78.

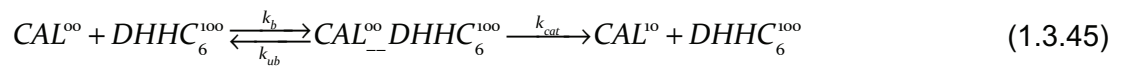
Since the network included two proteins that we already studied, we took advantage of their models and use them as starting point for the reconstruction. The first protein that has been implemented in the model was DHHC6, this step was particularly easy since we kept the structure of the model exactly as seen in Figure 5.5A. In this model DHHC16 can palmitoylate DHHC6 on three sites, and the reactions of palmitoylation and depalmitoylation are described using tQSSA kinetics. DHHC16 concentration is fixed and defines the maximum speed of modification (since $V_{\max} = k_{cat} \cdot [DHHC16]$). The only change made to the DHHC6 single protein model is the inclusion of a cycle of palmitoylation for APT2 as shown in equation (1.3.44).



In this model the concentration of APT2 is fixed, but the enzyme can undergo palmitoylation and depalmitoylation reactions, defined by simple mass action kinetics. Therefore, the amount of active APT2 can be tuned by changing the equilibrium constant of the cycle of APT2 palmitoylation/depalmitoylation.

For what concern calnexin we kept the topology used for the single protein model visible in Figure 4.1. The reaction mechanisms of calnexin modification required heavier modifications than what was done for DHHC6. The equations describing the model were modified to include explicitly the presence of DHHC6 in the model. DHHC6 have 8 different differentially palmitoylated species that can catalyse palmitoylation on calnexin, each one with different properties and efficiency. Each one the DHHC6 species can bind calnexin on one of the two sites available for

palmitoylation, catalyse the modification and release the protein. These steps of binding unbinding and catalysis were designed like a classic enzyme-substrate reaction in two steps, as visible in equation (1.3.45) for one DHHC6 configuration. Therefore all the equations of the enzymatic steps of calnexin modification described with tQSSA in the single protein model (see Table S4.2) were substituted by mass action terms describing binding, unbinding and catalysis of calnexin palmitoylation by DHHC6.



APT1 is modelled in the same way as APT2. APT1 have constant concentration but can undergo a cycle of palmitoylation/depalmitoylation that define the amount of active enzyme.

6.3 Preliminary results

6.3.1 Model calibration

Due to the complexity of the model, resulting from the high number of species and reactions, the procedure of parameter estimation (P.E.) resulted particularly challenging. Due to the increase in computation time required to simulate the model, we couldn't perform global optimization using genetic algorithm. The amount of time that is required to perform the optimization is prohibitive at the moment. Since we believe that global optimization has proven to be very valuable for P.E. we are currently working in different directions to speed up the computation and make the algorithm faster. At the present we already achieved a discrete gain in computation speed, but some more work needs to be done in order to reach the grade of efficiency required.

Even if a complete global P.E. has not yet been done, we managed to perform local optimization. This allowed us to obtain some sets of parameters that were used to check consistency with the previous models, and to perform some preliminary analysis. Moreover, we were still able to run global sensitivity analysis, since it only requires the specification of a parameter space in which to analyse the model.

After local optimization using least square fitting we obtained 20 different sets of parameters that we could use to simulate the model. To perform this kind of optimization in a reasonable amount many parameters were inherited from the models of DHHC6 and calnexin developed before. This helped during P.E. by reducing the number of parameters to estimate. Their value could be used as reference point from which to start the optimization, consistently speeding up the process. In particular, the values 44 parameters were taken from the sets of DHHC6 and calnexin model could be fixed during parameter estimation, while the value of new 60 parameters had to be estimated *de-novo*. The bounds on the 60 new parameters were define as done in section 3.3.3. Then we used the following procedure to generate their values:

1. First we selected the 20 best sets of parameters from the calnexin and DHHC6 model based on the ranking obtained from global optimization
2. The 11 parameters from the calnexin model and the 33 parameters from DHHC6 were then used to set the initial point for the local search, while the values of the remaining parameters were randomly generated in between the bound of each parameter. At the end we generated 20 starting points. Since the parameters of single models may not be optimal for the network model, the inherited parameter were allowed to vary $\pm 20\%$ during the optimization. The values of the remaining parameters were randomly generated in between the bound of each parameter.
3. We then optimized the 20 sets of parameters with least square fitting. After the optimization this set of parameters were used for model simulation and analysis. When we simulate the model we always repeat simulation for each set

of the 20 optimized sets, and we report as model output the median and the 1st and 3rd Quartile as a measure of the variance among the pool. The resulting parameters can be seen in Table 6.1.

This procedure allowed us to avoid estimation of all the parameters from scratch, reducing the calculation time consistently. While we work on reducing the computation time for global optimization we believe that the set estimated with this technique can be useful to check the model consistency and to generate some preliminary results

6.3.2 Global sensitivity analysis

To understand which parameters were more important in determining the different dynamics of the network we performed global sensitivity analysis. We were mainly interested in determine the important parameters on two main characteristics of the model:

1. Total amount of DHHC6/calnexin in steady state
2. The rate of palmitate incorporation

Since the model have 104 parameters, for an easier interpretation of the results and to reduce the computation time we grouped the parameters in 31 groups shown in. Table 6.2. For each analysis we used 128k samples and when we report the results we show only the sensitivity of those parameters which sum of the first order effect and total effect is bigger than 15% of the parameter with the highest sensitivity indexes. The bounds on the parameter space in which to perform S.A. were obtained from the set of optimized parameters. For each parameter we determined its maximum and minimum value among the sets, then the maximum value was doubled while the minimum was halved, obtaining the final values for the bounds on each parameter.

We first started to analyse the total protein in steady state, results are shown in Figure 6.3.

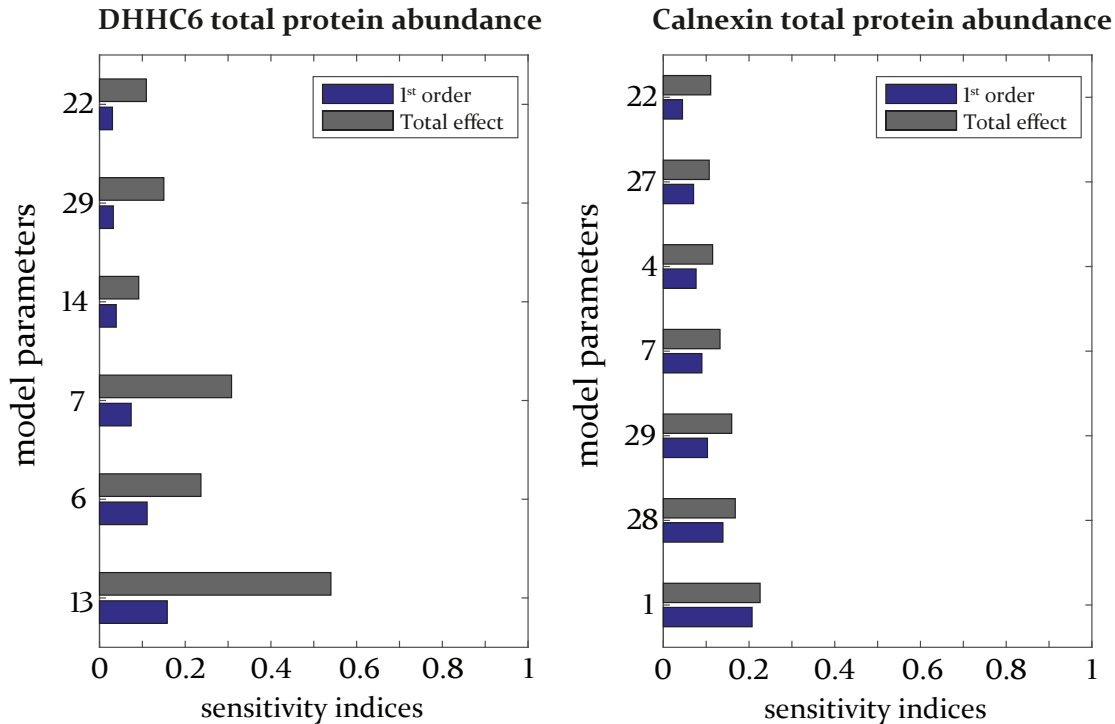


Figure 6.3. Global sensitivity analysis results for total DHHC6/calnexin in steady state.

Sensitivity analysis on the total abundance of DHHC6 and calnexin in steady state show interesting results. For what concern DHHC6 we can see that the 2 groups of parameter with the highest sensitivity are the trivial ones. DHHC6 is mostly influenced by the half-life of the DHHC6^{000} state and the synthesis/folding group. Since the bulk of DHHC6 lies in the unpalmitoylated state it seems logical that variation of the stability of this pool can have drastic effects on the amount of protein in the system. Similarly, it is reasonable that synthesis and folding play a major role in determining the abundance of DHHC6. Groups 7, 29 and 22 instead show how important is palmitoylation of DHHC6 to maintain its concentration. Group 7 represent the affinity towards DHHC16 and maximum speed of palmitoylation of the 1st site of DHHC6. In Chapter 5 - we have shown how differentially palmitoylated

DHHC6 can have very different stability, and that palmitoylation of the second and third site seems to have a stabilizing effect on the enzyme, especially if both the sites are modified at the same time. In fact, group 22, include parameters that define the maximum speeds of palmitoylation of multiple modifications, confirming that multiple palmitoylation events can greatly influence the abundance of the protein. Finally, group 29, which include only the concentration of DHHC16 in the model, highlight the importance of this enzyme in determining the abundance of DHHC6. Since DHHC16 is responsible for DHHC6 palmitoylation, which have drastic consequence on the enzyme stability, it seems reasonable that variation in its concentration can have dramatic effects on DHHC6 abundance.

For what concern calnexin, the first thing that we notice is that different parameters that affect DHHC6 concentration have a relevant effect on calnexin as well. In particular group 7, the parameters that determine the amount of palmitoylation on site 1 of DHHC6, have a marked effect on the abundance of calnexin. Since site 1 of DHHC6 confer the most activity when modified it is logical that group 7 have a primary role in determining the amount of palmitoylated calnexin, and as consequence its stability. The presence of group 22, the palmitoylation speeds of multiple modifications, confirm that the palmitoylation pattern of DHHC6 is fundamental in determining the abundance of calnexin. Interestingly one of the groups that have the most effect on calnexin concentration is group 29, or the concentration of DHHC16. The result is of particular interest because it correlates the concentration of calnexin with the one of DHHC16, suggesting that the palmitoylation cascade have an important regulatory mechanism that determines the abundance of palmitoylation substrates. Furthermore DHHC16 concentration affect more the abundance of calnexin than the one of DHHC6. This is in agreement with the results of Chapter 5, where we have shown that DHHC6 can redistribute itself in stable palmitoylation pattern, allowing the enzyme to mitigate the consequences of variation in DHHC16/APT2 concentration. Since this mechanism is not present in calnexin, this can explain why DHHC16 concentration affects more calnexin than DHHC6. The second most important group in determining calnexin concentration is group 16,

which includes all the catalytic rates of DHHC6. Because palmitoylation have a high stabilization effect on calnexin it is natural that the activity of DHHC6 plays a role in determining its concentration.

In Chapter 4 - and Chapter 5 - we studied the effects of palmitoylation on calnexin and DHHC6 respectively. A common outcome from the analysis of the models was the presence of a delay in palmitoylation of both proteins. In particular, we measured an average time for acquiring a first palmitoylation of 6 hours for DHHC6, and 8 hours for calnexin. Although we couldn't find an explanation for these delays, the fact that DHHC6 and calnexin belong to the same palmitoylation cascade and that the two delays are similar, with calnexin having a slightly bigger delay, let us thought that the two delays in palmitoylation are related. This suggest that these delays may be regulate by some regulatory mechanism in the upstream part of the palmitoylation cascade. Previous analysis indicates that the delay in palmitoylation observed is related with the speed of palmitate incorporation rate. In fact, increasing or decreasing the palmitate incorporation rate allow to tune the delay in palmitoylation. We therefore analysed which parameters are most influential on the palmitate incorporation rate of calnexin and DHHC6, to see if we could find some hint that could explain the delay observed in their modification. Results are shown in Figure 6.4.

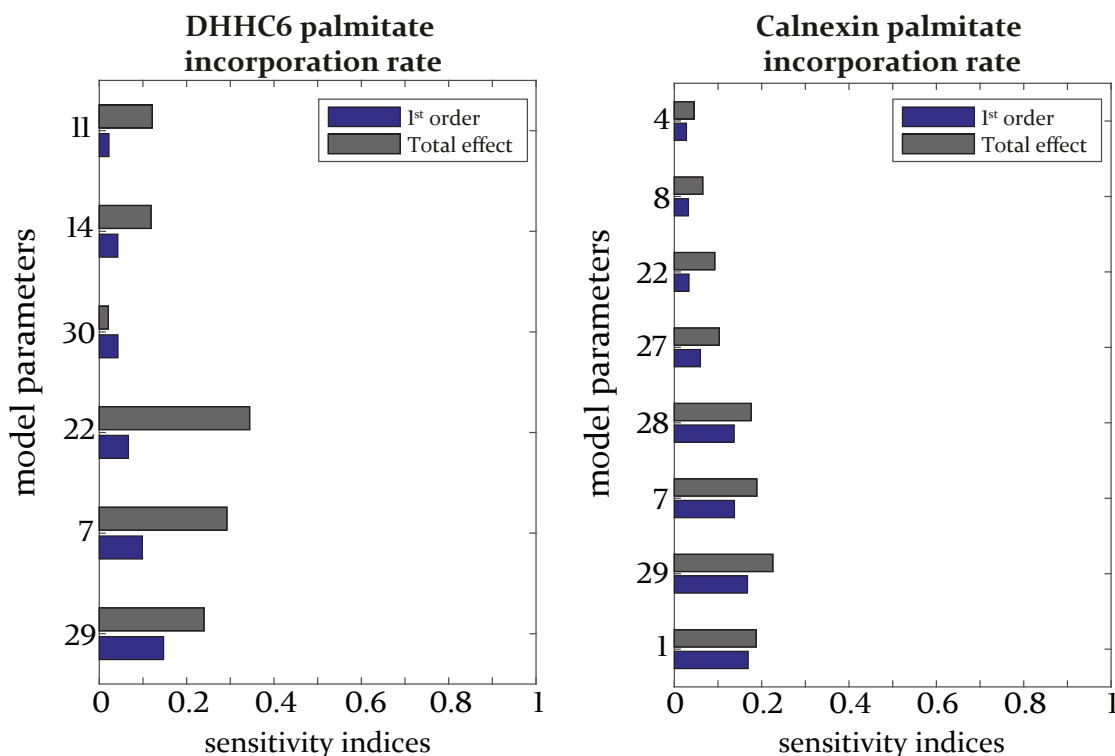


Figure 6.4. Global sensitivity analysis on palmitate incorporation rate of DHHC6 and calnexin.

Remarkably one of the first things that we can notice is that the two incorporation rates are somehow related. In fact 50% of the parameter groups that are most influential for DHHC6 are the same as in calnexin. In particular, the concentration of DHHC16 in the model is the most influential parameter for DHHC6 and the second one for calnexin, just after the synthesis and folding group. Therefore, the analysis shows that the amount of available DHHC16 is fundamental to determine the rate of palmitate incorporation both in DHHC6 and calnexin. As consequence, by tuning DHHC16 concentration it is possible to increase or decrease the delay in palmitoylation of both proteins. Of great importance for the palmitate rate incorporation is also the speed of palmitoylation of C^{100} and its affinity for DHHC16, for both DHHC6 and calnexin. This highlights the importance of the first site of DHHC6, which is the most palmitoylated. Moreover, being C^{100} the site that

contribute most to DHHC6 activity, it is also fundamental in determining the palmitate incorporation rate in calnexin.

6.4 Conclusion

In this chapter we developed a network model describing the palmitoylation cascade of calnexin. Through the use of rule-based modelling we described a system of 5 proteins that, although small, it was characterized by an elevated complexity. In fact, the network is composed by 705 different species and 2500 reactions fluxes. The aim of this model was to better characterize the regulatory effects of palmitoylation on DHHC6 and its substrate calnexin.

We used global sensitivity to analyse two main features of this network. First we determined which groups of parameters had the most impact on the concentration of DHHC6 and calnexin in steady state. Results suggest that palmitoylation is of primary importance in determining the concentration of both proteins. In fact, all the groups composed by parameters defining the palmitoylation speed and affinity of DHHC6 and calnexin show a high impact on their concentrations. A remarkable result is that DHHC16 affect both concentration of DHHC6 and calnexin. Surprisingly, DHHC16 have a stronger effect on calnexin abundance than on DHHC6. This may be partly explained by a mechanism observed for DHHC6 that allows the enzyme to redistribute itself in stable palmitoylation patterns upon variation in concentration of DHHC16/APT2, allowing the enzyme to mitigate the consequences of their variations. This mechanism was only observed in DHHC6 and this may explain why the effect of DHHC16 is stronger for calnexin.

We next investigated if there was correlation between the delay in palmitoylation observed in calnexin and DHHC6. Sensitivity analysis show that the delays strictly correlate with the abundance of DHHC16. Results suggest that the both delays can be tuned by regulating the abundance of DHHC16 in the model.

A part from the results presented in this chapter, we believe that a strong point of this model is that it is easy to include other proteins that are part of the palmitoylation network of DHHC6 or calnexin. Since rule-based representation is a very modular approach, different parts of the network with similar topology can be described by the same set of rules.

Appendix C

Table 6.1. Results of parameter estimation. Since the estimation gave us 20 different sets of parameters, here we report the median of the sets along with the 1st and 3rd quartile of their distribution as measure of variance.

Parameters	Median	1st quartile	3rd Quartile
ks_CAL	0.82	0.75	0.83
kf_CAL	2.32	2.28	2.53
kd_CAL_F	0.15	0.14	0.15
kd_CAL_C10	0.07	0.06	0.08
kd_CAL_C11	0.01	0.01	0.02
Vb_C	0.07	0.06	0.10
Kmb2c1_C	0.45	0.28	0.60
Kmb2c2_C	0.43	0.37	0.54
Kmb1c1_C	40.46	40.09	40.75
Kmb1c2_C	54.38	49.56	60.00
ks_DHHC6	0.00	0.00	0.00
kf_DHHC6	0.55	0.31	1.06
Vf_C100	41.20	20.60	67.24
Kmf1_C100	0.02	0.02	0.02
Vf_C010	0.00	0.00	0.01
Kmf2_C010	0.00	0.00	0.00
Vf_C001	0.39	0.19	1.43
Kmf3_C001	0.01	0.01	0.01
Vb_C001	0.01	0.00	0.01
Kmb3_C001	0.00	0.00	0.00
Vb_C010	0.06	0.04	0.08
Kmb2_C010	0.02	0.01	0.02
Vb_C100	0.10	0.06	0.13
Kmb1_C100	0.00	0.00	0.01
kd_DHHC6_C000	0.02	0.01	0.03
kd_DHHC6_C100	0.19	0.12	0.26
kd_DHHC6_C010	0.00	0.00	0.01
kd_DHHC6_C001	0.03	0.01	0.05
kd_DHHC6_C110	0.10	0.02	0.25
kd_DHHC6_C101	0.04	0.02	0.14
kd_DHHC6_C011	0.00	0.00	0.00
kd_DHHC6_C111	1.77	0.89	2.51
kdU_DHHC6	0.23	0.18	0.39
Vf_C110	0.07	0.04	0.13
Vf_C101	0.66	0.35	1.10

Vf_C011	2.78	1.23	4.43
Vf_C111	0.46	0.25	2.15
Vb_C110	10.47	2.67	39.28
Vb_C101	5.99	2.20	30.54
Vb_C011	0.54	0.31	1.06
Vb_C111	1.01	0.33	2.18
kb_F_C10_C100	574.12	444.09	2384.03
kb_F_C01_C100	1547.18	579.51	2382.11
kb_F_C10_C010	1772.97	454.68	2451.99
kb_F_C01_C010	1440.67	907.78	2116.88
kb_F_C10_C001	984.00	397.53	1975.93
kb_F_C01_C001	1565.76	701.71	2126.85
kb_F_C10_C110	1460.34	629.38	2040.15
kb_F_C01_C110	1461.17	716.10	1957.78
kb_F_C10_C101	1597.82	519.91	2016.46
kb_F_C01_C101	1610.47	994.28	2272.91
kb_F_C10_C011	1105.12	489.33	1753.46
kb_F_C01_C011	1273.73	677.72	2467.13
kb_F_C10_C111	1796.36	637.87	2372.16
kb_F_C01_C111	2090.47	1686.89	2675.30
kb_C10_C100	1796.06	511.35	2541.36
kb_C01_C100	1032.90	394.68	1617.20
kb_C10_C010	747.99	253.90	1514.79
kb_C01_C010	1186.21	374.62	1822.74
kb_C10_C001	1459.45	633.08	2114.70
kb_C01_C001	938.52	616.61	1919.97
kb_C10_C110	1609.10	746.71	2524.62
kb_C01_C110	1421.60	1283.14	1872.56
kb_C10_C101	1719.40	812.67	2299.10
kb_C01_C101	1450.46	546.37	2469.12
kb_C10_C011	1396.69	923.63	2158.86
kb_C01_C011	1424.86	1056.51	2098.47
kb_C10_C111	1265.07	880.96	2217.02
kb_C01_C111	1214.31	673.56	2346.14
kub_C11_C10_C100	14557.10	8238.98	25328.83
kub_C11_C01_C100	18076.52	10395.02	28162.32
kub_C11_C10_C010	18817.44	12353.80	23993.54
kub_C11_C01_C010	14865.94	6338.80	30618.20
kub_C11_C10_C001	18729.23	8754.34	23715.85
kub_C11_C01_C001	16340.85	7524.98	27739.78

kub_C11_C10_C110	13438.51	10602.11	21438.10
kub_C11_C01_C110	17659.27	10655.60	24615.26
kub_C11_C10_C101	16897.88	8878.68	19995.65
kub_C11_C01_C101	16461.87	11672.93	24252.05
kub_C11_C10_C011	15094.88	6705.71	27219.15
kub_C11_C01_C011	16779.43	6953.72	26753.14
kub_C11_C10_C111	16310.65	6765.18	24588.84
kub_C11_C01_C111	13333.91	7991.89	20640.97
kub_C10_C100	14259.61	9780.59	22849.53
kub_C01_C100	20286.20	6608.15	24526.87
kub_C10_C010	17267.99	9747.69	22459.91
kub_C01_C010	15282.03	8877.10	24046.52
kub_C10_C001	11568.80	7488.47	22912.61
kub_C01_C001	14028.39	7687.69	21388.72
kub_C10_C110	21101.87	12622.95	25774.21
kub_C01_C110	16747.18	9043.18	20764.60
kub_C10_C101	12832.81	9135.13	22773.56
kub_C01_C101	17838.67	8424.17	24212.11
kub_C10_C011	12016.82	5239.54	24600.16
kub_C01_C011	9684.21	5906.55	13916.32
kub_C10_C111	14342.10	10453.84	20385.06
kub_C01_C111	12614.34	7056.44	20884.17
kcat_C100	50.11	31.94	71.37
kcat_C010	22.86	9.40	64.79
kcat_C001	58.63	31.65	99.91
kcat_C110	106.53	66.23	235.60
kcat_C101	88.50	43.45	218.70
kcat_C011	158.25	66.04	309.79
kcat_C111	174.94	66.07	203.59
DHHC16Tot	0.15	0.09	0.25
APT1Tot	0.15	0.11	0.19
APT2Tot	0.67	0.61	0.89

Table 6.2. Groups used for global sensitivity analysis

Group	Parameters	Description
1	ks_CAL, kf_CAL	synthesis and folding of calnexin
2	kd_Cal_F	degradation of folded calnexin (not palmitoylated)
3	$kd_Cal_C10_C01$	degradation of calnexin C^{10} and C^{01}
4	kd_Cal_C11	degradation of calnexin C^{11}
5	$Vb_CAL, Kmb2cl_CAL, Kmb2cl_CAL, Kmb2cl_CAL, Kmb2cl_CAL$	Calnexin Maximum depalmitoylation speed and affinity for APTI
6	ks_DHHC6, kf_DHHC6	Synthesis and folding of DHHC6
7	Vf_C100_DHHC6, Kmf_C100	DHHC6 Maximum palmitoylation speed of site C^{100} and affinity for DHHC16
8	Vf_C010_DHHC6, Kmf_C010	DHHC6 Maximum palmitoylation speed of site C^{010} and affinity for DHHC16

9	<i>Vf_C001_DHHC6, Kmf_C001</i>	DHHC6 Maximum palmitoylation speed of site C ⁰⁰¹ and affinity for DHHC16
10	<i>Vb_C001_DHHC6, Kmb_C001</i>	DHHC6 Maximum depalmitoylation speed of site C ⁰⁰¹ and affinity for APT2
11	<i>Vb_C010_DHHC6, Kmb_C010</i>	DHHC6 Maximum depalmitoylation speed of site C ⁰¹⁰ and affinity for APT2
12	<i>Vb_C100_DHHC6, Kmb_C100</i>	DHHC6 Maximum depalmitoylation speed of site C ¹⁰⁰ and affinity for APT2
13	<i>kd_DHHC6_C000</i>	degradation rate of DHHC6 ⁰⁰⁰
14	<i>kd_DHHC6_C100</i>	degradation rate of DHHC6 ¹⁰⁰
15	<i>kd_DHHC6_C010</i>	degradation rate of DHHC6 ⁰¹⁰
16	<i>kd_DHHC6_C001</i>	degradation rate of DHHC6 ⁰⁰¹
17	<i>kd_DHHC6_C110</i>	degradation rate of DHHC6 ¹¹⁰
18	<i>kd_DHHC6_C101</i>	degradation rate of DHHC6 ¹⁰¹

19	kd_DHHC6_C0II	degradation rate of DHHC6 ^{0II}
20	kd_DHHC6_CIII	degradation rate of DHHC6 ^{III}
21	kd_DHHC6_U	degradation rate of DHHC6 unfolded
22	$Vf_C1I0_DHHC6, Vf_C10I_DHHC6, Vf_C0II_DHHC6, Vf_CIII_DHHC6$	Maximum palmitoylation speeds of multiple palmitoylated states
23	$Vb_C1I0_DHHC6, Vb_C10I_DHHC6, Vb_C0II_DHHC6, Vb_CIII_DHHC6$	Maximum depalmitoylation speeds of multiple palmitoylated states
24	$kb_F_C10_C100, kb_F_C0I_C100, kb_F_C10_C0I0, kb_F_C0I_C0I0, kb_F_C10_C00I, kb_F_C0I_C00I, kb_F_C10_C1I0, kb_F_C10_C10I, kb_F_C0I_C10I, kb_F_C10_C0II, kb_F_C0I_C0II, kb_F_C10_CIII, kb_F_C0I_CIII$	Binding rates of folded calnexin to all the possible palmitoylated states of DHHC6
25	$kb_C10_C100, kb_C0I_C100, kb_C10_C0I0, kb_C0I_C0I0, kb_C0I_C0I0, kb_C0I_C00I, kb_C0I_C00I, kb_C10_C1I0, kb_C0I_C1I0, kb_C10_C10I, kb_C0I_C10I, kb_C10_C0II, kb_C0I_C0II$	Binding rates of mono palmitoylated calnexin to all the possible palmitoylated states of DHHC6

	<i>,kb_C10_C111,kb_C01_C111</i>	
26	<i>kub_C11_C10_C100,kub_C11_C01_C100,kub_C11_C10_C010,kub_C11_C01_C010,kub_C11_C10_C110,kub_C11_C01_C01_C110,kub_C11_C10_C101,kub_C11_C01_C10_C011,kub_C11_C01_C01_C101,kub_C11_C10_C011,kub_C11_C01_C111,kub_C11_C01_C111</i>	Unbinding rates of double palmitoylated calnexin from all the possible palmitoylated states of DHHC6
27	<i>kub_C10_C100,kub_C01_C100,kub_C10_C010,kub_C01_C010,kub_C10_C001,kub_C01_C001,kub_C10_C110,kub_C01_C110,kub_C10_C101,kub_C01_C101,kub_C10_C101,kub_C01_C101,kub_C01_C111,kub_C01_C111</i>	unbinding rates of mono palmitoylated calnexin from all the possible palmitoylated states of DHHC6states
28	<i>kcat_C100,kcat_C010,kcat_C001,kcat_C110,kcat_C101,kcat_C011,kcat_C111</i>	catalytic rates of all the possible palmitoylated states of DHHC6
29	<i>DHHC16tot</i>	total concentration of DHHC16
30	<i>APT2tot</i>	total concentration of DHHC16
31	<i>APT1tot</i>	total concentration of DHHC16

Chapter 7 - Future work

With this project we revealed new insights about palmitoylation and its regulation. Despite this, there are still many questions about palmitoylation that are waiting to be answered. We believe that the models and the tools developed during this project will be very valuable in the future for the study of palmitoylation at the system level. In particular, the palmitoylation cascade developed in this project is easily expandable and will be useful for the analysis of other parts of the palmitoylation network, allowing the study of palmitoylation in different directions.

In the near future we plan to terminate the setting up of the palmitoylation network model. We believe that the model would greatly benefit from a global parameter estimation approach. Until now many parameters of the palmitoylation cascade were inherited from the models of calnexin and DHHC6, but some of them were estimated using a local optimization approach. A global approach would allow to estimate more parameter sets, increasing the ability of the model to reproduce the dynamics of the palmitoylation cascade. What we propose is approach similar to what has been done for calnexin and DHHC6, using genetic algorithm for parameter estimation. This task is currently under development; but the complexity of the model makes this kind of approach unfeasible for now, mainly because of the amount of time that this kind of optimization requires. We are currently working to optimize the procedure and make it faster. Global optimization would allow the network model to fit better the palmitoylation dynamics, allowing us to perform more accurate and reliable analysis improving the quality and precision of our findings. In particular, we would like to better characterize the regulatory effect of palmitoylation on the cascade, studying the effects of the different modifications of DHHC6 by DHHC16 on the palmitoylation dynamics of calnexin. It will also be interesting to observe how the system would react, and compensate, to perturbations of DHHC16 and APTs concentration and activity, since this components are fundamental in the regulation of the cycle of palmitoylation/depalmitoylation of DHHC6 and calnexin.

For what concern long term plans, we have set some goals from which we believe the network palmitoylation model would greatly benefit. One of the best achievements of this project was the development of a robust mechanism of reaction describing palmitoylation, along with an efficient technique for palmitoylation network reconstruction and the tools for its analysis. Despite the considerable number of features implemented in the model, there are still different elements that can be added to increase its accuracy and prediction capabilities. If we want to improve the design of the model we think that it would be beneficial to start working in three main directions.

The last model considered in this project represent only a very small subset of the cellular palmitoylation networks. This model has been used to study the properties of a palmitoylation cascade, a mechanism that seems to appear frequently in the palmitoylation network. Although useful, we think that more proteins need to be included in the model to start capturing the dynamics of palmitoylation at the cellular level. An intuitive way to expand the model would be to include other DHHC6 substrates, focusing on those for which experimental results are available. A good starting point would be for example Climp63, a type II transmembrane protein resident in the ER, but also present on the plasma membrane. This protein is palmitoylated on one site from DHHC6 and DHHC2, and it is not excluded that other palmitoyltransferases may contribute to its modification. Although the effects of palmitoylation are not fully understood it seems that this modification regulates the shuttling of Climp63 from the ER to the plasma membrane and contribute also to the formation of dimers. This protein would be ideal for the inclusion in the palmitoylation network since it is being actively studied in our laboratory. Therefore, different experiments useful for model calibration are already available. The addition of other proteins to the network would allow to study the regulation of palmitoylation of multiple substrates. It would be also interesting to investigate if the different palmitoylation state of DHHC6 have the same activity across all its substrates. Moreover expanding the model in this direction will probably help to understand how a very low abundant protein like DHHC6 is able to palmitoylate different substrates

that are very abundant in the cell, including some of the most abundant proteins in the cell like calnexin. In fact, proteome studies have estimated that calnexin is at least 300 times more abundant than DHHC6.

Another fundamental direction in which to expand our network model would be to start the modelling of other DHHC enzymes. There are 23 DHHC enzymes in human, and the most part of them, 16 identified so far, are themselves palmitoylated. This implies that DHHC enzymes are interacting with each other, forming a regulatory network that can be understood/investigated only under a network which integrates all DHHC enzymes and the interactions between them. An obvious choice from which to start would be to add to the model those DHHCs that have already been proven to be palmitoylated. Since these enzymes are being modified, it is probable that these DHHCs and all their substrates follow a very similar palmitoylation cascade to what is observed for DHHC6. This would allow to re-use parts of the network model that have been already developed to describe other parts of the cellular palmitoylation network. If this is the case, the work done until now would greatly facilitate and speed up the reconstruction of the palmitoylation cascade of those enzymes along with their substrates. The reason is that rule-based modelling, the approach that we have used to design the network of calnexin palmitoylation, is highly modular and allows easy re-utilization of reaction mechanics. We believe that expansion of the network in these two directions would be too difficult from the point of view of modelling due to the possibility of re-using what has been already done to describe new parts of the network.

One aspect that need to be carefully considered is that the expansion of the model that we are proposing would not be possible without experimental assays that would fuel the model with data for its calibration and validation. Experiments would be necessary to identify the actors in the palmitoylation cascade of palmitoylated DHHCs, and data on the dynamics of their synthesis, palmitoylation and localization would be necessary to correctly train the model. It is useful to remind that due to the lack of information on the kinetics of DHHC enzymes and their substrates, a

consistent part of this project implied the use of experiments to investigate the palmitoylation network. Therefore, an even broader experimental approach would be needed as the project advance. Of particular usefulness would be experiments to identify palmitoylation and interactions among DHHCs, along with data on their kinetics of synthesis, palmitoylation and localizations, as it has been done for DHHC6 and calnexin.

The reaction mechanism proposed by Goldbeter and Koshland, have proven to be able to describe the enzymatic mechanism of protein palmitoylation. This reaction kinetic has been used along with rule-based modelling to design the first network of palmitoylation, which described the palmitoylation cascade of the ER chaperone calnexin. These two techniques combined together allowed us to greatly improve our knowledge of palmitoylation, but the model can still be improved, to increase its accuracy, usefulness and prediction capabilities. One important improvement that we want to propose would be to explicitly include the pool of palmitate in the dynamics of the model. In this work we have always assumed that palmitate was present in excess amount, and therefore it was never a rate-limiting substrate in the reactions of palmitoylation. It is known that palmitate is available as palmitoyl-CoA at nanomolar concentration in the cell. It is still not known if palmitate is enriched in specific microdomains or if its concentration is more or less constant in the cell. Implementing the palmitate pool in the model would help to understand in which condition palmitate could become a rate-limiting step in the signalling network of palmitoylation. Moreover, it would be possible to study the consequences of palmitate starvation on the capabilities of the signalling network. The implementation of the pool of palmitate in the model would require data from experimental techniques that can measure the abundance of palmitate in the different compartments of the cell, and which would be able to point out where the palmitate is enriched. These experiments would be very beneficial to the model, allowing an increased level of detail that may reveal additional information on palmitate dynamics.

Another implementation that would be very valuable for the network model would be the explicit inclusion of cellular compartments. This feature would be extremely useful since different studies have shown that palmitoylation have a major role in regulation of protein localization among the cell. Palmitoylation regulates the localization of different soluble and membrane proteins, and this is valid also for DHHCs enzymes as well. Implementing compartments would be useful to analyse the distribution of differentially palmitoylated proteins, and how palmitoylation rearrange those proteins in case of a stress event. Since most part of DHHCs are localized between the ER and the Golgi apparatus, and many membrane proteins have been shown to shuttle between the membranes of those compartments and the plasma membrane, modelling these three compartments would be a good starting point. All the proteins considered in this project, (including some proteins that are not present in this thesis) reside in one of these three membranes, with some being able to shuttle between them. Data about distribution of different palmitoylated proteins based on their palmitoylation status would be valuable for an initial calibration of the model, which would then be used to analyse the capabilities of the network on other palmitoylation substrates.

Chapter 8 - Conclusions

With this work we have contributed to increase the knowledge about palmitoylation and its regulatory mechanisms and we have set up the basis for the reconstruction of the different cellular palmitoylation networks. These goals are of primary importance to uncover the role of palmitoylation as regulatory mechanism. This project required the characterization of palmitoylation as enzymatic reaction, the development of protein palmitoylation models for validation of the proposed reaction mechanism and a reconstruction of a palmitoylation sub-network, also used for validation of the proposed network topology. A considerable amount of time has been dedicated in the development of tools that will speed up the computational time required for model simulation and investigation.

In order to achieve the different objectives we have followed a bottom up approach. The first step was to characterize the enzymatic reaction of palmitoylation, and then to translate it into a reaction mechanism capable of capture and reproduce correctly the dynamics of protein palmitoylation/depalmitoylation (Chapter 3 -). The choice has fallen on the well-established mechanism of reversible protein modification described by Goldbeter and Koshland in [72]. This mechanism of reaction has been successfully used to describe signalling networks in hundreds of publications, since it works particularly well in the case of signalling cascades and pathways. The Goldbeter and Koshland system describing MAPK protein cascade phosphorylation is modelled with mass action terms. In the case of the palmitoylation systems studied here and taking into account the fact how little these systems have been studied in the past we decided to reduce the complexity by deriving Michaelian-type expressions for the reaction rates of the network. The validity of the approximation for in-vivo reactions, often characterized by the presence of equal amount of enzyme and substrate, was

ensured by using a version of the approximation called tQSSA which overcomes the inherent inability of the Michaelis-Menten (MM) assumption of capturing the properties of such systems.

Once defined the reaction mechanisms the next challenge was to verify their capabilities in describing reactions of palmitoylation of different substrates. In Chapter 4 - we show the design and implementation of the first mathematical model of protein palmitoylation, describing the modification process of the ER chaperone calnexin. Results of model analysis clearly show that the chosen reaction mechanism accurately describes palmitoylation/depalmitoylation dynamics, showing good agreement with experimental data during model calibration and validation. Even more convincing is the fact that a number of model predictions have been confirmed through experiments, adding to the validation of the tQSSA mechanism. Both the model and the experimental approach used in this project provided unprecedented understanding of the palmitoylation process. With the calnexin model we manage to characterize the steady state distribution of this protein, showing that calnexin is mainly double palmitoylated. This is because the two sites show cooperativity and therefore they are sequentially palmitoylated. Interestingly we were able to tune the amount of palmitoylated calnexin from 0 to 100% just by modulating the activity of DHHC6. The model helped to quantify the stabilization effect of palmitoylation on calnexin that was first observed with experiments. Moreover, the model revealed something that was never observed with experiments, in fact. we observed that palmitoylation of calnexin doesn't happen just after its synthesis, but in average it takes 8hrs for a molecule of calnexin to get double palmitoylated All these results are discussed in details in section 4.5.

To further validate our studies, the reaction mechanism of palmitoylation has been tested on DHHC6 as well (Chapter 5 -). DHHC6 is a transmembrane enzyme that catalyses palmitoylation of calnexin, its modification is essential for being active. Following an approach similar to what was done for calnexin, we have built a kinetic model of DHHC6 palmitoylation. As for calnexin, the model has been able to

accurately describe DHHC6 synthesis and palmitoylation dynamics. First we predicted the steady state distribution of DHHC6 in the different palmitoylation states. Remarkably, we found a situation completely opposite to what was observed for calnexin. In fact, DHHC6 is mainly non palmitoylated in steady state, suggesting that the majority of the enzyme is not active. Analysis of the stability upon palmitoylation of each site show two opposite effects. While palmitoylation of site 1 cause a huge drop in half-life, modification of site 2 and 3 stabilize the enzyme increasing its half-life. This suggest that the stability of DHHC6 is tightly regulated by palmitoylation. Analysis of model fluxes revealed consistent palmitoylation/depalmitoylation cycles between C^{000} and C^{100} , suggesting that site 1 is preferentially palmitoylated with respect to the other sites. Upon overexpression of DHHC16, or partial APT silencing, we observed a redistribution of the majority of the enzyme in the C^{011} state, with the majority of the cycles now happening between C^{100} and C^{011} . The same analysis was repeated on CAA mutant showing that when site 2 and 3 are missing we lose 70% of the enzyme due to the instability caused by over palmitoylation of site 1. Results suggest that the role of site 2 and 3 is to allow DHHC6 to tolerate variation in the concentration of DHHC16 or APT2 avoiding enzyme depletion. Stochastic simulations have shown that, as for calnexin, there is a delay between DHHC6 synthesis and palmitoylation. We measured an average palmitoylation time of 6 hrs for DHHC6, slightly shorter than what observed for calnexin. This suggests that the two delays are related, with the delay in palmitoylation observed for calnexin being dependent by the one of DHHC6. It is probable that the delays observed are caused by an upstream regulatory mechanism in the palmitoylation cascade of calnexin. All the results about DHHC6 analysis are described in detail in section 5.3.

In the last part of the thesis we studied palmitoylation at the system level with the aim to get some insights in the regulatory mechanisms of cellular palmitoylation networks. For this purpose we decided to reconstruct the palmitoylation cascade of calnexin (Chapter 6 -). A significant amount of time has been invested in the design of the sub-network model. Due to the complexity of the model we had to find an automated method to generate ordinary differential equations (ODEs). To solve the issue we used

a particular kind of modelling technique called “Rule-based modelling” [140,141]. This kind of modelling allowed to specify a set of rules that defined how reactions take place and which species are involved in a particular reaction. The rule-set can then be translated automatically into a set of ODEs. This approach hugely simplified the design of the model since the rule-set is much simpler than the model that it describes. This is especially true for palmitoylation cascades, which are normally made of repeated reaction schemes. It was therefore possible to use the same set of rules to describe different parts of the model, tremendously simplifying the workload. Because of the complexity of this model, a considerable amount of time has also been invested in optimizing the model and speeding up the computational time for the simulations. At the time of writing of my thesis the optimization process is still ongoing, therefore we were only able to generate preliminary results from a subset of model parameters derived from our previous works. The time to execute a complete parameter estimation with the network model requires still a considerable amount of time that is prohibitive with the current equipment, but big reductions in computational time have already been obtained. Even if we analysed only a subset of parameters, the model proved to be useful to generate new hypotheses on the functioning of the palmitoylation network that were not intuitive from single protein models. Results from global sensitivity revealed that palmitoylation is the main determinant in regulating the abundance of DHHC6 and calnexin in the cell. Moreover, the delay observed in DHHC6 and calnexin palmitoylation depend mainly from the amount of DHHC16 available to palmitoylated DHHC6. By tuning DHHC16 concentration it is possible to increase or decrease the delay in palmitoylation of both proteins. Results of the study of the calnexin palmitoylation cascade are described in detail in section 6.4.

The computational approach used in this project has proven to be an extremely useful tool for the study of palmitoylation networks. The power of this approach is that with mathematical modelling we can organize experiments in a consistent mathematical model. The model can then be analysed to gain further understanding that is hidden in the complexity of the system. The modelling approach have also been useful to drive experiments and to suggest modifications of the experimental protocols that led

to an increased understanding of the experimental techniques and of palmitoylation. But the usefulness of this approach becomes more evident when it is combined in cycles of experiment/modelling. While the experiments fuel the model and allow its calibration and validation, subsequent analysis of the model can generate new hypotheses that can then be tested experimentally, allowing for further comprehension of palmitoylation. The methods and the models developed in this project have been designed keeping in mind the final objective of reconstruction of the palmitoylation networks of the human cell. Therefore, the design of the network, the tool for its analysis and the methods for model setting construction have the potential to be applied to any part of the palmitoylation network, making them useful for the future studies on palmitoylation.

Chapter 9 - References

- [1] Knorre DG, Kudryashova NV, Godovikova TS. Chemical and Functional Aspects of Posttranslational Modification of Proteins. *Acta Naturae* 2009;1:29–51.
- [2] Beltrao P, Bork P, Krogan NJ, van Noort V. Evolution and functional cross-talk of protein post-translational modifications. *Mol Syst Biol* 2013;9. doi:10.1002/msb.201304521.
- [3] Humphrey SJ, James DE, Mann M. Protein Phosphorylation: A Major Switch Mechanism for Metabolic Regulation. *Trends Endocrinol Metab* n.d.;26:676–87. doi:10.1016/j.tem.2015.09.013.
- [4] Ciechanover A. The unravelling of the ubiquitin system. *Nat Rev Mol Cell Biol* 2015;16:322–4. doi:10.1038/nrm3982.
- [5] Burnett G, Kennedy EP. THE ENZYMATIC PHOSPHORYLATION OF PROTEINS. *J Biol Chem* 1954;211:969–80.
- [6] Cohen P. The regulation of protein function by multisite phosphorylation – a 25 year update. *Trends Biochem Sci* 2000;25:596–601. doi:10.1016/S0968-0004(00)01712-6.
- [7] Pinna LA, Ruzzene M. How do protein kinases recognize their substrates? *Biochim Biophys Acta BBA - Mol Cell Res* 1996;1314:191–225. doi:10.1016/S0167-4889(96)00083-3.
- [8] Kolarova M, García-Sierra F, Bartos A, Ricny J, Ripova D. Structure and Pathology of Tau Protein in Alzheimer Disease. *Int J Alzheimers Dis* 2012;2012:731526. doi:10.1155/2012/731526.
- [9] Stewart T, Sossi V, Aasly JO, Wszolek ZK, Uitti RJ, Hasegawa K, et al. Phosphorylated α -synuclein in Parkinson's disease: correlation depends on disease severity. *Acta Neuropathol Commun* 2015;3:1–11. doi:10.1186/s40478-015-0185-3.
- [10] Maiuri MC, Kroemer G. Essential Role for Oxidative Phosphorylation in Cancer Progression. *Cell Metab* n.d.;21:11–2. doi:10.1016/j.cmet.2014.12.013.
- [11] Goldstein G, Scheid M, Hammerling U, Schlesinger DH, Niall HD, Boyse EA. Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells. *Proc Natl Acad Sci U S A* 1975;72:11–5.
- [12] Glickman MH, Ciechanover A. The Ubiquitin-Proteasome Proteolytic Pathway: Destruction for the Sake of Construction. *Physiol Rev* 2002;82:373–428. doi:10.1152/physrev.00027.2001.
- [13] Schnell JD, Hicke L. Non-traditional Functions of Ubiquitin and Ubiquitin-binding Proteins. *J Biol Chem* 2003;278:35857–60. doi:10.1074/jbc.R300018200.
- [14] Mukhopadhyay D, Riezman H. Proteasome-Independent Functions of Ubiquitin in Endocytosis and Signaling. *Science* 2007;315:201–5. doi:10.1126/science.1127085.
- [15] Blaskovic S, Blanc M, van der Goot FG. What does S-palmitoylation do to membrane proteins? *FEBS J* 2013;280:2766–74. doi:10.1111/febs.12263.
- [16] Blanc M, David F, Abrami L, Migliozi D, Armand F, Børgi J, et al. SwissPalm: Protein Palmitoylation database [version 1; referees: 3 approved]. *F1000Research* 2015;4. doi:10.12688/f1000research.6464.1.

- [17] Aicart-Ramos C, Valero RA, Rodriguez-Crespo I. Protein palmitoylation and subcellular trafficking. *Biochim Biophys Acta BBA - Biomembr* 2011;1808:2981–94. doi:10.1016/j.bbamem.2011.07.009.
- [18] Martin BR, Cravatt BF. Large-scale profiling of protein palmitoylation in mammalian cells. *Nat Meth* 2009;6:135–8. doi:10.1038/nmeth.1293.
- [19] The UniProt Consortium. UniProt: a hub for protein information. *Nucleic Acids Res* 2015;43:D204–12. doi:10.1093/nar/gku989.
- [20] Salaun C, Greaves J, Chamberlain L. The intracellular dynamic of protein palmitoylation. *J Cell Biol* 2010;191:1229–38. doi:10.1083/jcb.201008160.
- [21] Merrick BA, Dhungana S, Williams JG, Aloor JJ, Peddada S, Tomer KB, et al. Proteomic Profiling of S-acylated Macrophage Proteins Identifies a Role for Palmitoylation in Mitochondrial Targeting of Phospholipid Scramblase 3. *Mol Cell Proteomics* 2011;10.
- [22] Sharma C, Rabinovitz I, Hemler ME. Palmitoylation by DHHC3 is critical for the function, expression, and stability of integrin $\alpha 6\beta 4$. *Cell Mol Life Sci* 2012;69:2233–44. doi:10.1007/s00018-012-0924-6.
- [23] Zheng H, Pearsall EA, Hurst DP, Zhang Y, Chu J, Zhou Y, et al. Palmitoylation and membrane cholesterol stabilize μ -opioid receptor homodimerization and G protein coupling. *BMC Cell Biol* 2012;13:1–18. doi:10.1186/1471-2121-13-6.
- [24] Huang K, Kang MH, Askew C, Kang R, Sanders SS, Wan J, et al. Palmitoylation and function of glial glutamate transporter-1 is reduced in the YAC128 mouse model of Huntington disease. *Spec Issue Non-Mamm Models Neuropsychiatr Dis* 2010;40:207–15. doi:10.1016/j.nbd.2010.05.027.
- [25] Tsuchiya Y, Akashi M, Matsuda M, Goto K, Miyata Y, Node K, et al. Involvement of the Protein Kinase CK2 in the Regulation of Mammalian Circadian Rhythms. *Sci Signal* 2009;2:ra26-. doi:10.1126/scisignal.2000305.
- [26] Lee H-S, Hwang CY, Shin S-Y, Kwon K-S, Cho K-H. MLK3 Is Part of a Feedback Mechanism That Regulates Different Cellular Responses to Reactive Oxygen Species. *Sci Signal* 2014;7:ra52-. doi:10.1126/scisignal.2005260.
- [27] Lai AZ, Cory S, Zhao H, Gigoux M, Monast A, Guiot M-C, et al. Dynamic Reprogramming of Signaling Upon Met Inhibition Reveals a Mechanism of Drug Resistance in Gastric Cancer. *Sci Signal* 2014;7:ra38-. doi:10.1126/scisignal.2004839.
- [28] Pauli E-K, Chan YK, Davis ME, Gableske S, Wang MK, Feister KF, et al. The Ubiquitin-Specific Protease USP15 Promotes RIG-I-Mediated Antiviral Signaling by Deubiquitylating TRIM25. *Sci Signal* 2014;7:ra3-. doi:10.1126/scisignal.2004577.
- [29] Rocks O, Peyker A, Kahms M, Verveer PJ, Koerner C, Lumbierres M, et al. An acylation cycle regulates localization and activity of palmitoylated Ras isoforms. *Science* 2005;307:1746–52. doi:10.1126/science.1105654.
- [30] Tarrant MK, Cole PA. The Chemical Biology of Protein Phosphorylation. *Annu Rev Biochem* 2009;78:797–825. doi:10.1146/annurev.biochem.78.070907.103047.
- [31] Johnson LN. The regulation of protein phosphorylation. *Biochem Soc Trans* 2009;37:627–41. doi:10.1042/BST0370627.
- [32] Fukata Y, Fukata M. Protein palmitoylation in neuronal development and synaptic plasticity. *Nat Rev Neurosci* 2010;11:161–75. doi:10.1038/nrn2788.
- [33] Resh MD. Covalent lipid modifications of proteins. *Curr Biol* n.d.;23:R431–5. doi:10.1016/j.cub.2013.04.024.

- [34] Lakkaraju AK, Abrami L, Lemmin T, Blaskovic S, Kunz B, Kihara A, et al. Palmitoylated calnexin is a key component of the ribosome–translocon complex. *EMBO J* 2012;31:1823–35. doi:10.1038/emboj.2012.15.
- [35] Braun V, Rehn K. Chemical Characterization, Spatial Distribution and Function of a Lipoprotein (Murein-Lipoprotein) of the *E. coli* Cell Wall. *Eur J Biochem* 1969;10:426–38. doi:10.1111/j.1432-1033.1969.tb00707.x.
- [36] Braun V, Hantke K, Henning U. Characterization of the free form of murein-lipoprotein from the outer membrane of *Escherichia coli* B/r. *FEBS Lett* 1975;60:26–8. doi:10.1016/0014-5793(75)80410-8.
- [37] Magee AI, Gutierrez L, McKay IA, Marshall CJ, Hall A. Dynamic fatty acylation of p21N-ras. *EMBO J* 1987;6:3353–7.
- [38] Degtyarev MY, Spiegel AM, Jones TL. Increased palmitoylation of the Gs protein alpha subunit after activation by the beta-adrenergic receptor or cholera toxin. *J Biol Chem* 1993;268:23769–72.
- [39] Dallavilla T, Abrami L, Sandoz PA, Savoglidis G, Hatzimanikatis V, van der Goot FG. Model-Driven Understanding of Palmitoylation Dynamics: Regulated Acylation of the Endoplasmic Reticulum Chaperone Calnexin. *PLoS Comput Biol* 2016;12:e1004774. doi:10.1371/journal.pcbi.1004774.
- [40] Yount JS, Moltedo B, Yang Y-Y, Charron G, Moran TM, López CB, et al. Palmitoylome profiling reveals S-palmitoylation–dependent antiviral activity of IFITM3. *Nat Chem Biol* 2010;6:610–4. doi:10.1038/nchembio.405.
- [41] Yang W, Di Vizio D, Kirchner M, Steen H, Freeman MR. Proteome Scale Characterization of Human S-Acylated Proteins in Lipid Raft-enriched and Non-raft Membranes. *Mol Cell Proteomics* 2010;9:54–70. doi:10.1074/mcp.M800448-MCP200.
- [42] Abrami L, Kunz B, Iacovache I, van der Goot FG. Palmitoylation and ubiquitination regulate exit of the Wnt signaling protein LRP6 from the endoplasmic reticulum. *Proc Natl Acad Sci* 2008;105:5384–9. doi:10.1073/pnas.0710389105.
- [43] Flannery AR, Czibener C, Andrews NW. Palmitoylation-dependent association with CD63 targets the Ca²⁺ sensor synaptotagmin VII to lysosomes. *J Cell Biol* 2010;191:599–613. doi:10.1083/jcb.201003021.
- [44] Resh MD. Fatty acylation of proteins: new insights into membrane targeting of myristoylated and palmitoylated proteins. *Biochim Biophys Acta BBA - Mol Cell Res* 1999;1451:1–16. doi:http://dx.doi.org/10.1016/S0167-4889(99)00075-0.
- [45] Maeda A, Okano K, Park PS-H, Lem J, Crouch RK, Maeda T, et al. Palmitoylation stabilizes unliganded rod opsin. *Proc Natl Acad Sci* 2010;107:8428–33. doi:10.1073/pnas.1000640107.
- [46] Abrami L, Leppla SH, van der Goot FG. Receptor palmitoylation and ubiquitination regulate anthrax toxin endocytosis. *J Cell Biol* 2006;172:309–20. doi:10.1083/jcb.200507067.
- [47] Blaskovic S, Adibekian A, Blanc M, van der Goot GF. Mechanistic effects of protein palmitoylation and the cellular consequences thereof. *Chem Phys Lipids* 2014;180:44–52. doi:10.1016/j.chemphyslip.2014.02.001.
- [48] Hancock JF, Paterson H, Marshall CJ. A polybasic domain or palmitoylation is required in addition to the CAAX motif to localize p21ras to the plasma membrane. *Cell* 1990;63:133–9. doi:10.1016/0092-8674(90)90294-O.
- [49] Parenti M, Viganó MA, Newman CM, Milligan G, Magee AI. A novel N-terminal motif for palmitoylation of G-protein alpha subunits. *Biochem J* 1993;291:349–53.

- [50] Linder ME, Middleton P, Hepler JR, Taussig R, Gilman AG, Mumby SM. Lipid modifications of G proteins: alpha subunits are palmitoylated. *Proc Natl Acad Sci* 1993;90:3675–9. doi:10.1073/pnas.90.8.3675.
- [51] Joseph M, Nagaraj R. Interaction of Peptides Corresponding to Fatty Acylation Sites in Proteins with Model Membranes. *J Biol Chem* 1995;270:16749–55. doi:10.1074/jbc.270.28.16749.
- [52] Charollais J, Van Der Goot FG. Palmitoylation of membrane proteins (Review). *Mol Membr Biol* 2009;26:55–66. doi:10.1080/09687680802620369.
- [53] Mitchell DA, Vasudevan A, Linder ME, Deschenes RJ. Thematic review series: Lipid Posttranslational Modifications. Protein palmitoylation by a family of DHHC protein S-acyltransferases. *J Lipid Res* 2006;47:1118–27. doi:10.1194/jlr.R600007-JLR200.
- [54] Lobo S, Greentree WK, Linder ME, Deschenes RJ. Identification of a Ras Palmitoyltransferase in *Saccharomyces cerevisiae*. *J Biol Chem* 2002;277:41268–73. doi:10.1074/jbc.M206573200.
- [55] Roth AF, Feng Y, Chen L, Davis NG. The yeast DHHC cysteine-rich domain protein Akr1p is a palmitoyl transferase. *J Cell Biol* 2002;159:23–8. doi:10.1083/jcb.200206120.
- [56] Ohno Y, Kihara A, Sano T, Igarashi Y. Intracellular localization and tissue-specific distribution of human and yeast DHHC cysteine-rich domain-containing proteins. *Biochim Biophys Acta* 2006;1761:474–83. doi:S1388-1981(06)00070-9 [pii] 10.1016/j.bbaliip.2006.03.010.
- [57] Zhou F, Xue Y, Yao X, Xu Y. CSS-Palm: palmitoylation site prediction with a clustering and scoring strategy (CSS). *Bioinformatics* 2006;22:894–6. doi:10.1093/bioinformatics/btl013.
- [58] Lakkaraju A, Abrami L, Lemmin T, Blaskovic S, Kunz B, Kihara A, et al. Palmitoylated calnexin is a key component of the ribosome-translocon complex. *EMBO J* 2012;31:1823–35. doi:10.1038/emboj.2012.15.
- [59] Huang K, Sanders S, Singaraja R, Orban P, Cijssouw T, Arstikaitis P, et al. Neuronal palmitoyl acyl transferases exhibit distinct substrate specificity. *FASEB J* 2009;23:2605–15. doi:10.1096/fj.08-127399.
- [60] Greaves J, Carmichael JA, Chamberlain LH. The palmitoyl transferase DHHC2 targets a dynamic membrane cycling pathway: Regulation by a C-terminal domain. *Mol Biol Cell* 2011;22:1887–95. doi:10.1091/mbc.E10-11-0924.
- [61] Greaves J, Gorleku OA, Salaun C, Chamberlain LH. Palmitoylation of the SNAP25 Protein Family: SPECIFICITY AND REGULATION BY DHHC PALMITOYL TRANSFERASES. *J Biol Chem* 2010;285:24629–38. doi:10.1074/jbc.M110.119289.
- [62] Tsutsumi R, Fukata Y, Noritake J, Iwanaga T, Perez F, Fukata M. Identification of G Protein α Subunit-Palmitoylating Enzyme. *Mol Cell Biol* 2009;29:435–47. doi:10.1128/MCB.01144-08.
- [63] Korycka J, Łach A, Heger E, Bogusławska DM, Wolny M, Toporkiewicz M, et al. Human DHHC proteins: A spotlight on the hidden player of palmitoylation. *Eur J Cell Biol* 2012;91:107–17. doi:10.1016/j.ejcb.2011.09.013.
- [64] Swarthout JT, Lobo S, Farh L, Croke MR, Greentree WK, Deschenes RJ, et al. DHHC9 and GCP16 Constitute a Human Protein Fatty Acyltransferase with Specificity for H- and N-Ras. *J Biol Chem* 2005;280:31141–8. doi:10.1074/jbc.M504113200.
- [65] Zeidman R, Jackson CS, Magee AI. Protein acyl thioesterases (Review). *Mol Membr Biol* 2009;26:32–41. doi:10.1080/09687680802629329.

- [66] Chamberlain LH, Shipston MJ. The Physiology of Protein S-acylation. *Physiol Rev* 2015;95:341–76. doi:10.1152/physrev.00032.2014.
- [67] Duncan JA, Gilman AG. A Cytoplasmic Acyl-Protein Thioesterase That Removes Palmitate from G Protein α Subunits and p21RAS. *J Biol Chem* 1998;273:15830–7. doi:10.1074/jbc.273.25.15830.
- [68] Tomatis VM, Trenchi A, Gomez GA, Daniotti JL. Acyl-Protein Thioesterase 2 Catalyzes the Deacylation of Peripheral Membrane-Associated GAP-43. *PLoS ONE* 2010;5:e15045. doi:10.1371/journal.pone.0015045.
- [69] Kolch W, Calder M, Gilbert D. When kinases meet mathematics: the systems biology of MAPK signalling. *FEBS Lett* 2005;579:1891–5. doi:10.1016/j.febslet.2005.02.002.
- [70] Vaudry D, Stork PJS, Lazarovici P, Eiden LE. Signaling Pathways for PC12 Cell Differentiation: Making the Right Connections. *Science* 2002;296:1648. doi:10.1126/science.1071552.
- [71] Ferrell JE, Machleder EM. <div xmlns="http://www.w3.org/1999/xhtml">The Biochemical Basis of an All-or-None Cell Fate Switch in Xenopus Oocytes</div>. *Science* 1998;280:895. doi:10.1126/science.280.5365.895.
- [72] Goldbeter A, Koshland DE. An amplified sensitivity arising from covalent modification in biological systems. *Proc Natl Acad Sci U S A* 1981;78:6840–4.
- [73] Beck M, Schmidt A, Malmstroem J, Claassen M, Ori A, Szymborska A, et al. The quantitative proteome of a human cell line. *Mol Syst Biol* 2011;7. doi:10.1038/msb.2011.82.
- [74] Nagaraj N, Wisniewski JR, Geiger T, Cox J, Kircher M, Kelso J, et al. Deep proteome and transcriptome mapping of a human cancer cell line. *Mol Syst Biol* 2011;7:548–548. doi:10.1038/msb.2011.81.
- [75] Pedersen MG, Bersani AM, Bersani E, Cortese G. The total quasi-steady-state approximation for complex enzyme reactions. 5th Vienna Int Conf Math Model Sci Comput Electron Eng 2006 Int Conf Comput Sci Dyn Syst Comput Asp 2008;79:1010–9. doi:10.1016/j.matcom.2008.02.009.
- [76] Pedersen MG, Bersani AM, Bersani E. Quasi steady-state approximations in complex intracellular signal transduction networks – a word of caution. *J Math Chem* 2008;43:1318–44. doi:10.1007/s10910-007-9248-4.
- [77] Tzafirri AR. Michaelis-Menten kinetics at high enzyme concentrations. *Bull Math Biol* 2003;65:1111–29. doi:10.1016/S0092-8240(03)00059-4.
- [78] Borghans JAM, de Boer RJ, Segel LA. Extending the quasi-steady state approximation by changing variables. *Bull Math Biol* 1996;58:43–63. doi:10.1007/BF02458281.
- [79] Man K-F, Tang K-S, Kwong S. Genetic algorithms: concepts and applications. *IEEE Trans Ind Electron* 1996;43:519–34.
- [80] Deb K, Kalyanmoy D. *Multi-Objective Optimization Using Evolutionary Algorithms*. New York, NY, USA: John Wiley & Sons, Inc.; 2001.
- [81] Saltelli A, Ratto M, Andres T, Campolongo F, Cariboni J, Gatelli D, et al. *Front Matter. Glob. Sensit. Anal. Primer*, John Wiley & Sons, Ltd; 2007, p. i–xi.
- [82] Iooss B, Lemaître P. A Review on Global Sensitivity Analysis Methods. In: Dellino G, Meloni C, editors. *Uncertain. Manag. Simul.-Optim. Complex Syst. Algorithms Appl.*, Boston, MA: Springer US; 2015, p. 101–22.
- [83] Sobol' IM, Tarantola S, Gatelli D, Kucherenko SS, Mauntz W. Estimating the approximation error when fixing unessential factors in global sensitivity analysis. *Reliab Eng Syst Saf* 2007;92:957–60. doi:10.1016/j.ress.2006.07.001.

- [84] Sanft KR, Gillespie DT, Petzold LR. Legitimacy of the Stochastic Michaelis-menten Approximation. *IET Syst Biol* 2011;5:58–69.
- [85] Rao CV, Arkin AP. Stochastic chemical kinetics and the quasi-steady-state assumption: Application to the Gillespie algorithm. *J Chem Phys* 2003;118:4999–5010. doi:http://dx.doi.org/10.1063/1.1545446.
- [86] MacNamara S, Bersani AM, Burrage K, Sidje RB. Stochastic chemical kinetics and the total quasi-steady-state assumption: Application to the stochastic simulation algorithm and chemical master equation. *J Chem Phys* 2008;129. doi:http://dx.doi.org/10.1063/1.2971036.
- [87] Gillespie DT. Exact stochastic simulation of coupled chemical reactions. *J Phys Chem* 1977;81:2340–61. doi:10.1021/j100540a008.
- [88] Helenius A, Aebi M. Roles of N-linked glycans in the endoplasmic reticulum. *Annu Rev Biochem* 2004;73:1019–49.
- [89] Lynes E, Raturi A, Shenkman M, Ortiz Sandoval C, Yap M, Wu J, et al. Palmitoylation is the switch that assigns calnexin to quality control or ER Ca²⁺ signaling. *J Cell Sci* 2013;126:3893–903. doi:10.1242/jcs.125856.
- [90] Lakkaraju AK, van der Goot F. Calnexin controls the STAT3-mediated transcriptional response to EGF. *Mol Cell* 2013;51:386–96. doi:10.1016/j.molcel.2013.07.009.
- [91] Lynes EM, Bui M, Yap MC, Benson MD, Schneider B, Ellgaard L, et al. Palmitoylated TMX and calnexin target to the mitochondria-associated membrane. *Embo J* 2011;doi: 10.1038/emboj.2011.384. doi:10.1038/emboj.2011.384.
- [92] Chevet E, Smirle J, Cameron P, Thomas D, Bergeron J. Calnexin phosphorylation: linking cytoplasmic signalling to endoplasmic reticulum luminal functions. *Semin Cell Dev Biol* 2010;21:486–90. doi:10.1016/j.semcdb.2009.12.005.
- [93] Chevet E, Wong H, Gerber D, Cochet C, Fazel A, Cameron P, et al. Phosphorylation by CK2 and MAPK enhances calnexin association with ribosomes. *EMBO J* 1999;18:3655–66. doi:10.1093/emboj/18.13.3655.
- [94] Takizawa T, Tatematsu C, Watanabe K, Kato K, Nakanishi Y. Cleavage of calnexin caused by apoptotic stimuli: implication for the regulation of apoptosis. *J Biochem (Tokyo)* 2004;136:399–405. doi:10.1093/jb/mvh133.
- [95] Orton R, Sturm O, Vyshemirsky V, Calder M, Gilbert D, Kolch W. Computational modelling of the receptor-tyrosine-kinase-activated MAPK pathway. *Biochem J* 2005;392:249–61. doi:10.1042/BJ20050908.
- [96] Bhalla U, Iyengar R. Emergent properties of networks of biological signaling pathways. *Science* 1999;283:381–7. doi:10.1126/science.283.5400.381.
- [97] Dosztanyi Z, Csizmok V, Tompa P, Simon I. IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* 2005;21:3433–4. doi:10.1093/bioinformatics/bti541.
- [98] Markevich NI, Hoek JB, Kholodenko BN. Signaling switches and bistability arising from multisite phosphorylation in protein kinase cascades. *J Cell Biol* 2004;164:353–9. doi:10.1083/jcb.200308060.
- [99] Segel LA. On the validity of the steady state assumption of enzyme kinetics. *Bull Math Biol* 1988;50:579–93.
- [100] Thomson M, Gunawardena J. The rational parameterisation theorem for multisite post-translational modification systems. *J Theor Biol* 2009;261:626–36. doi:10.1016/j.jtbi.2009.09.003.

- [101] Foster LJ, de Hoog CL, Zhang Y, Zhang Y, Xie X, Mootha VK, et al. A mammalian organelle map by protein correlation profiling. *Cell* 2006;125:187–99. doi:10.1016/j.cell.2006.03.022.
- [102] Gillespie DT. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J Comput Phys* 1976;22. doi:10.1016/0021-9991(76)90041-3.
- [103] Korkhov VM, Milan-Lobo L, Zuber B, Farhan H, Schmid JA, Freissmuth M, et al. Peptide-based interactions with calnexin target misassembled membrane proteins into endoplasmic reticulum-derived multilamellar bodies. *J Mol Biol* 2008;378:337–52. doi:10.1016/j.jmb.2008.02.056.
- [104] Klopfenstein D, Klumperman J, Lustig A, Kammerer R, Oorschot V, Hauri H. Subdomain-specific localization of CLIMP-63 (p63) in the endoplasmic reticulum is mediated by its luminal alpha-helical segment. *J Cell Biol* 2001;153:1287–300. doi:10.1083/jcb.153.6.1287.
- [105] Lee D, Kraus A, Prins D, Groenendyk J, Aubry I, Liu WX, et al. UBC9-dependent Association between Calnexin and Protein Tyrosine Phosphatase 1B (PTP1B) at the Endoplasmic Reticulum. *J Biol Chem* 2015;290:5725–38. doi:10.1074/jbc.M114.635474.
- [106] Fredericks GJ, Hoffmann FW, Rose AH, Osterheld HJ, Hess FM, Mercier F, et al. Stable expression and function of the inositol 1,4,5-triphosphate receptor requires palmitoylation by a DHHc6/selenoprotein K complex. *Proc Natl Acad Sci U S A* 2014;111:16478–83. doi:10.1073/pnas.1417176111.
- [107] Du S, Zhou J, Jia Y, Huang K. SelK is a novel ER stress-regulated protein and protects HepG2 cells from ER stress agent-induced apoptosis. *Arch Biochem Biophys* 2010;502:137–43. doi:10.1016/j.abb.2010.08.001.
- [108] Vartak N, Papke B, Grecco HE, Rossmannek L, Waldmann H, Hedberg C, et al. The autodepalmitoylating activity of APT maintains the spatial organization of palmitoylated membrane proteins. *Biophys J* 2014;106:93–105. doi:10.1016/j.bpj.2013.11.024.
- [109] Kong E, Peng S, Chandra G, Sarkar C, Zhang Z, Bagh MB, et al. Dynamic palmitoylation links cytosol-membrane shuttling of acyl-protein thioesterase-1 and acyl-protein thioesterase-2 with that of proto-oncogene H-ras product and growth-associated protein-43. *J Biol Chem* 2013;288:9112–25. doi:10.1074/jbc.M112.421073.
- [110] Kang M, Day CA, Kenworthy AK, DiBenedetto E. Simplified equation to extract diffusion coefficients from confocal FRAP data. *Traffic* 2012;13:1589–600. doi:10.1111/tra.12008.
- [111] Axelrod D, Koppel DE, Schlessinger J, Elson E, Webb WW. Mobility measurement by analysis of fluorescence photobleaching recovery kinetics. *Biophys J* 1976;16:1055–69. doi:10.1016/S0006-3495(76)85755-4.
- [112] O.Palsson B. *Systems Biology Simulation of Dynamic Network States*. 2011.
- [113] Sanft KR, Gillespie DT, Petzold LR. Legitimacy of the stochastic Michaelis-Menten approximation. *IET Syst Biol* 2011;5:58. doi:10.1049/iet-syb.2009.0057.
- [114] Macnamara S, Bersani A, Burrage K, Sidje R. Stochastic chemical kinetics and the total quasi-steady-state assumption: application to the stochastic simulation algorithm and chemical master equation. *J Chem Phys* 2008;129. doi:10.1063/1.2971036.

- [115] Yokoi N, Fukata Y, Sekiya A, Murakami T, Kobayashi K, Fukata M. Identification of PSD-95 Depalmitoylating Enzymes. *J Neurosci* 2016;36:6431–44. doi:10.1523/JNEUROSCI.0419-16.2016.
- [116] Chamberlain LH, Lemonidis K, Sanchez-Perez M, Werno MW, Gorleku OA, Greaves J. Palmitoylation and the trafficking of peripheral membrane proteins. *Biochem Soc Trans* 2013;41:62–66. doi:10.1042/BST20120243.
- [117] Conibear E, Davis NG. Palmitoylation and depalmitoylation dynamics at a glance. *J Cell Sci* 2010;123:4007–4010. doi:10.1242/jcs.059287.
- [118] Montersino A, Thomas GM. Slippery signaling: Palmitoylation-dependent control of neuronal kinase localization and activity. *Mol Membr Biol* 2015;32:179–88. doi:10.1080/09687688.2016.1182652.
- [119] Wu R, Pruitt Z, Runkle M, Scerif G, Aslin RN. A neural signature of rapid category-based target selection as a function of intra-item perceptual similarity, despite inter-item dissimilarity. *Atten Percept Psychophys* 2016;78:749–60. doi:10.3758/s13414-015-1039-6.
- [120] Thuma F, Heiler S, Schnölzer M, Zöller M. Palmitoylated claudin7 captured in glycolipid-enriched membrane microdomains promotes metastasis via associated transmembrane and cytosolic molecules. *Oncotarget* 2016. doi:10.18632/oncotarget.8928.
- [121] Fairbank M, Huang K, El-Husseini A, Nabi IR. RING finger palmitoylation of the endoplasmic reticulum Gp78 E3 ubiquitin ligase. *FEBS Lett* 2012;586:2488–93. doi:10.1016/j.febslet.2012.06.011.
- [122] Fredericks GJ, Hoffmann FW, Rose AH, Osterheld HJ, Hess FM, Mercier F, et al. Stable expression and function of the inositol 1,4,5-triphosphate receptor requires palmitoylation by a DHHC6/selenoprotein K complex. *Proc Natl Acad Sci* 2014;111:16478–83. doi:10.1073/pnas.1417176111.
- [123] Senyilmaz D, Virtue S, Xu X, Tan CY, Griffin JL, Miller AK, et al. Regulation of mitochondrial morphology and function by stearylolation of TFR1. *Nature* 2015;525:124–8.
- [124] Gorleku OA, Barns A-M, Prescott GR, Greaves J, Chamberlain LH. Endoplasmic Reticulum Localization of DHHC Palmitoyltransferases Mediated by Lysine-based Sorting Signals. *J Biol Chem* 2011;286:39573–84. doi:10.1074/jbc.M111.272369.
- [125] Li Y, Martin BR, Cravatt BF, Hofmann SL. DHHC5 Protein Palmitoylates Flotillin-2 and Is Rapidly Degraded on Induction of Neuronal Differentiation in Cultured Cells. *J Biol Chem* 2012;287:523–30. doi:10.1074/jbc.M111.306183.
- [126] Sandoz PA, van der Goot FG. How many lives does CLIMP-63 have? *Biochem Soc Trans* 2015;43:222–8. doi:10.1042/BST20140272.
- [127] Yang W, Di Vizio D, Kirchner M, Steen H, Freeman M. Proteome scale characterization of human S-acylated proteins in lipid raft-enriched and non-raft membranes. *Mol Cell Proteomics MCP* 2010;9:54–70. doi:10.1074/mcp.M800448-MCP200.
- [128] Rossin A, Durivault J, Chakhtoura-Feghali T, Lounnas N, Gagnoux-Palacios L, Hueber A-O. Fas palmitoylation by the palmitoyl acyltransferase DHHC7 regulates Fas stability. *Cell Death Differ* 2015;22:643–53.
- [129] Lai J, Linder ME. Oligomerization of DHHC Protein S-Acyltransferases. *J Biol Chem* 2013;288:22862–70. doi:10.1074/jbc.M113.458794.
- [130] Percher A, Ramakrishnan S, Thinon E, Yuan X, Yount JS, Hang HC. Mass-tag labeling reveals site-specific and endogenous levels of protein S-fatty

- acylation. Proc Natl Acad Sci 2016;113:4302–7. doi:10.1073/pnas.1602244113.
- [131] Werno MW, Chamberlain LH. S-acylation of the Insulin-Responsive Aminopeptidase (IRAP): Quantitative analysis and Identification of Modified Cysteines. *Sci Rep* 2015;5:12413.
- [132] Goldbeter A, Koshland D. An amplified sensitivity arising from covalent modification in biological systems. *Proc Natl Acad Sci U S A* 1981;78:6840–4. doi:10.1073/pnas.78.11.6840.
- [133] Pedersen MG, Bersani AM, Bersani E, Cortese G. The total quasi-steady-state approximation for complex enzyme reactions. *Math Comput Simul* 2008;79. doi:10.1016/j.matcom.2008.02.009.
- [134] Pedersen MG, Bersani AM, Bersani E. Quasi steady-state approximations in complex intracellular signal transduction networks - a word of caution. *J Math Chem* 2008;43:1318–44. doi:10.1007/S10910-007-9248-4.
- [135] Borghans J, De R, Segel L. Extending the quasi-steady state approximation by changing variables 1996. doi:10.1007/BF02458281.
- [136] Yount JS, Moltedo B, Yang Y-Y, Charron G, Moran TM, López CB, et al. Palmitoylome profiling reveals S-palmitoylation-dependent antiviral activity of IFITM3. *Nat Chem Biol* 2010;6:610–4. doi:10.1038/nchembio.405.
- [137] Merrick BA, Dhungana S, Williams JG, Aloor JJ, Peddada S, Tomer KB, et al. Proteomic Profiling of S-acylated Macrophage Proteins Identifies a Role for Palmitoylation in Mitochondrial Targeting of Phospholipid Scramblase 3. *Mol Cell Proteomics* 2011;10.
- [138] Beck M, Schmidt A, Malmstroem J, Claassen M, Ori A, Szymborska A, et al. The quantitative proteome of a human cell line. *Mol Syst Biol* 2011;7. doi:10.1038/msb.2011.82.
- [139] Nagaraj N, Wisniewski JR, Geiger T, Cox J, Kircher M, Kelso J, et al. Deep proteome and transcriptome mapping of a human cancer cell line. *Mol Syst Biol* 2011;7:548–548. doi:10.1038/msb.2011.81.
- [140] Faeder JR, Blinov ML, Hlavacek WS. Rule-Based Modeling of Biochemical Systems with BioNetGen. In: Maly VI, editor. *Syst. Biol.*, Totowa, NJ: Humana Press; 2009, p. 113–67.
- [141] Lily A Chylek and Leonard A Harris and James R Faeder and William S Hlavacek. Modeling for (physical) biologists: an introduction to the rule-based approach. *Phys Biol* 2015;12:45007.

Tiziano Dallavilla



📍 Avenue de Tivoli 32, Lausanne, CH-1007, Vaud, Switzerland

📞 +393282176230

☎ +41788508201

✉ tiziano.dallavilla@gmail.com

Profile

- PhD in Biotechnology and Bioengineering, BSc and MSc in Biotechnology with a major focus in bioinformatics.
- Interdisciplinary PhD research expertise: Biology, computational kinetic modelling, Systems biology, Biochemistry.
- High level teamwork skills: 4 years of working experience as interface between two leading laboratories.

Skills

Computational

- Operating systems: Linux/Unix/MacOS/Windows
- Programming languages: R, Java, Python, SQL, Matlab, C++, PERL
- Management and interrogation of database with SQL
- High performance parallel computing (SLURM)
- Analyses of large datasets
- Use of version control systems


Communication

- Brainstorm, flashlight
- Scientific communication (oral presentations, poster presentations)
- Scientific writing (reports, papers, grant applications)
- Productive teamwork

Modelling

- Reconstruction of cellular networks with kinetic modelling
- Techniques for model analysis (i.e. Global sensitivity analysis, Global parameter estimation, ANOVA, *in silico* labelling...)
- Large scale data analysis and visualization

Experience

 **Nov 2012-present** Research assistant, Laboratory of Computational Systems Biotechnology and Laboratory of Cell and Membrane Biology, EPFL (*Expected graduation: October 2016*).

Major deliverables

- Study of palmitoylation with the use of kinetic models
- Development of mathematical tools for model analysis
- Reconstruction of the first human palmitoylation network
- *In silico* prediction of protein properties affected by palmitoylation
- Cross-validation of the model-based predictions with *in vivo* experiments
- Development of semi-automated modelling framework for the reconstruction of the entire human palmitoylation network

Added value


- Responsible for the informatics services in the Laboratory of Computational Systems Biotechnology:
 - Management and maintenance of 3 servers for data storage
 - Set up and maintenance of version control system (GIT) for all the projects of the laboratory
 - Set up and maintenance of electronic equipment in the laboratory
- Teaching assistant for a total of 509 hours for Systems Biology and Biochemical Engineering courses

2002-2012 Part time salesman at food commerce family business

Education


PhD degree in Biotechnology and Bioengineering

École Polytechnique Fédérale de Lausanne (EPFL)

 Nov 2012-Oct 2016 (*Expected*)


M.S. in Bioinformatics

Università degli studi di Milano-Bicocca

 2011-2012

B.S. in Industrial and environmental Biotechnology

Università degli studi di Milano

 2006-2010

Languages

• Italian ● ● ● ● ●
(Mother tongue)

• English ● ● ● ● ●
(Professional working proficiency)

• French ● ○ ○ ○ ○
(Beginner)

Scientific publications

1. **Model-Driven Understanding of Palmitoylation Dynamics: Regulated Acylation of the Endoplasmic Reticulum Chaperone Calnexin**
Dallavilla T, Abrami L, Sandoz PA, Savoglidis G, Hatzimanikatis V, et al. (2016). PLoS Comput Biol 12(2): e1004774. doi: 10.1371/journal.pcbi.1004774.
2. **Function and dynamics of the first palmitoylation cascade DHHCl6-mediated control of DHHC6 activity: dynamics of a palmitoylation cascade.** Abrami L, Dallavilla T, Sandoz P, Demir M., Kunz B., Savoglidis G., Knott G., Guex N., Hatzimanikatis V. van der Goot F.G. (*Manuscript in preparation*).
3. **Study of palmitoylation networks with kinetic modelling.** Tiziano Dallavilla, *PhD thesis*.

Extracurricular activities

- Passion for informatics with a particular interest in hardware and assembly of high performance computers.
- Sports: snowboard freeride, mtb downhill, longboard, swimming.
- Passionate traveler and explorer, with interest towards Asian, South American and African countries.

