HAL
open science

DEVELOPMENT OF FUNCTIONAL PROTEOMICS APPROACHES FOR STUDYING RETROGRADE TRANSPORT

Shi Getao

To cite this version:

Shi Getao. DEVELOPMENT OF FUNCTIONAL PROTEOMICS APPROACHES FOR STUDYING RETROGRADE TRANSPORT. Biochemistry [q-bio.BM]. Ecole Normale Supérieure de Paris - ENS Paris, 2011. English. NNT : . tel-00635924

HAL Id: tel-00635924

https://theses.hal.science/tel-00635924v1

Submitted on 26 Oct 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Thèse de Doctorat

En vue de l'obtention du grade de

DOCTEUR
DE L'ÉCOLE NORMALE SUPÉRIEURE

École doctorale: Interdisciplinaire Européenne Frontières du Vivant

Discipline ou spécialité : Biologie et Chimie

Présentée et soutenue par : **Getao SHI**

le 17 Octobre 2011

Titre:

DEVELOPMENT OF FUNCTIONAL PROTEOMICS APPROACHES FOR STUDYING RETROGRADE TRANSPORT

Directeurs de thèse: JOHANNES Ludger et FLORENT Jean-Claude

Membres du jury (y compris les directeurs de thèse)

M. CHARNAY Patrick	Présidente
Mme. SAGAN Sandrine	Rapporteur
M. MELNYK Oleg	Rapporteur
M. LORD Mike	Examinateur

Numéro identifiant de la Thèse :

SUMMARY

Retrograde trafficking allows for proteins transport from the plasma membrane to the endoplasmic reticulum, via the Golgi apparatus, bypassing the degrading environment of lysosomes. Previous studies have shown that retrograde transport is crucial for cellular functions such as signaling, ion homeostasis, and the establishment of wnt morphogen gradients. Furthermore, retrograde transport plays an essential role in cellular uptake of pathogenic factors like protein toxins and virus proteins. However, the current list of cargo proteins is limited, and it is likely that many cellular functions of retrograde transport still remain to be discovered. Clearly, a strong need exists for further characterization of this transport route.

In this study, four different proteomics approaches have been developed aiming to identify membrane proteins taking retrograde transport route: SNAP-tag, sulfation, FKBP, and streptavidin. Amongst these, the SNAP-tag approach turned out to be the most efficient strategy to identify cargo candidates of the retrograde route. This strategy is based on the covalent modification of the plasma membrane proteome with a benzylguanine (BG) derivative. Only BG-tagged plasma membrane proteins that are then transported via the retrograde route can covalently couple to a TGN-localized SNAP-tag fusion protein. The approach was validated step-by-step using a well studied retrograde cargo protein, Shiga toxin B-subunit (STxB). We could show that STxB can be captured and identified by the SNAP-tag approach. Furthermore, coupled with LC-MS, new protein candidates of the retrograde route have been identified.

The other three approaches have guided our choice towards the most efficient proteomics strategy, by bringing up experimental opportunities and threats in the fields of organic chemistry and cell biology. Altogether, this thesis work has allowed to develop a novel proteomics strategy that could be applied to identify new cargo

candidates of the retrograde route. We have pioneered a dynamic proteomics approach that complements traditional proteomics. Moreover, the concept of using chemical tools to study retrograde transport can also be applied to other endocytic pathways.

Keywords:

Retrograde transport, Proteomics, SNAP-tag, Golgi apparatus

RÉSUMÉ FRANÇAIS

Le trafic rétrograde permet le transport des protéines de la membrane plasmique vers le réticulum endoplasmique, via l'appareil de Golgi, évité la dégradation des lysosomes. Des études antérieures ont montré que le transport rétrograde est crucial pour les fonctions cellulaires telles que la signalisation, l'homéostasie ionique, et l'établissement de gradients du morphogène Wnt. Par ailleurs, le transport rétrograde joue un rôle essentiel dans l'internalisation cellulaire des facteurs pathogènes comme les toxines protéiques et les protéines de virus. Toutefois, la liste actuelle des protéines cargos est limitée, et il est probable que de nombreuses fonctions cellulaires du transport rétrograde restent encore à découvrir. De toute évidence, un fort besoin existe pour une caractérisation plus poussée de cette voie de transport.

Dans cette étude, quatre différentes approches protéomiques ont été développées visant à identifier les protéines membranaires prenant la route du transport rétrograde: SNAP-tag, sulfatation, la FKBP, et la streptavidine. Parmi ceux-ci, l'approche SNAP-tag s'est avéré être la stratégie la plus efficace pour identifier les candidats du fret de la voie rétrograde. Cette stratégie est basée sur la modification covalente du protéome de la membrane plasmique avec un benzylguanine (BG) dérivés. Seules les protéines membranaires BG-taggées qui sont ensuite transportés par voie rétrograde peuvent coupler covalentement à une protéine de fusion de SNAP-tag localisée dans la TGN. L'approche a été validée, étape par étape, en utilisant une protéine cargo rétrograde bien étudiée, toxine Shiga sous-unité B (STxB). Nous avons pu montrer que les STxB peuvent être capturés et identifiées par l'approche SNAP-tag. Par ailleurs, couplé à la LC-MS, les candidats des nouvelle protéines de la voie rétrograde ont été identifiés.

Les trois autres approches ont guidé notre choix vers la stratégie la plus efficace en protéomique, en apportant des possibilités d'expérimentation et de défauts dans les

domaines de la chimie organique et en biologie cellulaire. Généralement, ce travail de thèse a permis de développer une nouvelle stratégie protéomique qui pourraient être appliquée pour identifier les nouvelles candidats de la route rétrograde. Nous avons mis au point une approche dynamique en protéomique qui complète la protéomique traditionnelle. Par ailleurs, le concept d'utiliser des outils chimiques pour étudier le transport rétrograde peut également être appliqué à d'autres voies d'endocytose.

Mots clés: Transport rétrograde, la protéomique, SNAP-tag, l'appareil de Golgi

To Chun’er CHEN(陈春儿)

To Xinfu SHI(施鑫富)

TABLE OF CONTENTS

TABLE OF CONTENTS **7**
TABLE OF FIGURES **10**
TABLE OF TABLES **12**
TABLE OF SCHEMES **13**
ABBREVATIONS **14**
INTRODUCTION **16**
1. Intracellular transport pathways 17
1.1 Biosynthetic/secretory pathway 19
1.2 Endocytic pathways 22
1.2.1 Late endocytic/degradation pathway 22
1.2.2 Recycling pathway 24
1.2.3 Retrograde transport 24
2. Retrograde transport 27
2.1 The functions of retrograde transport 27
2.2 Pathology 30
2.2.1 Shiga Toxin 30
2.2.2 Ricin 32
2.2.3 Cholera toxin 32
2.2.4 Virus 33
2.2.5 Other diseases 33
2.3 Retrograde transport machinery 34
2.3.1 Clathrin coats 37
2.3.2 Retromer 39
2.4 Tools to study retrograde transport 40
3. Proteomics 42
3.1 The functions and methods of proteomics 43
3.1.1 Identification and analysis of proteins 43
3.1.2 Differential-display proteomics 46
a). The two-dimensional gel approach 46
b). Quantitative comparison by mass spectrometry 46
c). Protein chips 47
3.1.3 Protein–protein interactions 48
a). Purification of protein complexes 48
b). Yeast two-hybrid screening 49
c). Phage display 51
3.2 Proteomics in cell biology 51
3.2.1. A mitochondrial protein complex that links apoptosis and glycolysis 51
3.2.2. New component in clathrin coated vesicles 52
3.2.3. Kinase pathways that regulate sex-specific functions in Plasmodium 54
4. Reaction schemes for proteomics approaches 55

4.1 SNAP-tag/AGT, a DNA repair enzyme 55
4.2 Sulfation 59
4.3 Rapamycin, FK506 and FKBP 63
4.4 Biotin and streptavidin 65
4.5 N-hydroxysuccinimide in protein coupling 68
RESULTS 70
1. SNAP-tag approach 71
1.1 Synthesis of the BG derivative 73
1.2 Setting-up the experimental system 73
1.2.1 Expressing SNAP-tag in the Golgi apparatus 73
a). TGN38 and SNAP-tag hybrid 75
b). SNAP-tag fusion protein with GalT 77
1.2.2 Functional validation 83
a). STxB-BG endocytosis assay 83
b). LC-MS detection 87
c). Antibody uptake assay 89
d). Cell permeability test 91
e). Cell surface modification 94
1.3 Proteomics screening 96
2. Sulfation approach 100
2.1 The synthesis of bSuPeRs 103
2.1.1 bSuPeR-3b synthesis 103
2.1.2 Click reaction 103
2.1.3 bSuPeR-carbamate synthesis 105
2.2 Experimental setup 105
2.2.1 Estimation of the efficiency for cell surface modification 105
2.2.2 bSuPeR-carbamate characterization 106
2.2.3 An important element for efficient cell surface modification: reaction pH 108
2.2.4 A key element for efficient cell surface modification: the peptide charge 108
3. FKBP approach 114
3.1 FK506-NHS synthesis 114
3.2 Verification strategy 116
3.2.1 In vitro binding 116
a). STxB coupling to FK506-NHS 116
b). Binding assay in solution 116
c). Binding assay on cell surface 118
3.2.2 Localization of FKBP12 to the Golgi apparatus 118
3.2.3 Functional validation 120
4. Streptavidin approach 120
MATERIALS AND METHODS 123
1. Chemistry part 124
2. Biology part. 136

DISCUSSION **141**
1. Identification of cargo candidates of the retrograde route 142
1.1 Integrin 142
1.2 Transferrin receptor 143
1.3 Ion transporters 143
1.4 Critical discussion 144
2. STxB as a model to verify proteomics approaches 145
3. Innovative proteomic approaches 146
3.1 Proteomics strategy for retrograde transport 146
3.2 Comparison with traditional proteomics approaches 147
3.3 Comparison of the four approaches 148
4. SNAP-tag for protein-protein interaction 150
5. Chemical probe synthesis 150
6 Negative charge and polyarginine 151
PERSPECTIVE **152**
ACKNOWLEDGEMENTS **155**
REFERENCES **156**
ANNEX **170**

TABLE OF FIGURES

Figure 1. The different pathways of intracellular vesicle trafficking

Figure 2. Models for intra-Golgi trafficking

Figure 3. Endocytosis pathways into cells

Figure 4. Recycling pathways

Figure 5. Scheme for retrograde transport

Figure 6. The structure of shiga toxin

Figure 7. The location of different elements in retrograde transport

Figure 8. The structure of clathrin and the position of clathrin triskelions in a clathrin coat

Figure 9. A strategy for mass spectrometric identification of proteins

Figure 10. Scheme for yeast two-hybrid screening

Figure 11. AGT suicide reaction

Figure 12. Coupling reaction between AGT and BG

Figure 13. AGT fusion protein labeling and different BG derivatives

Figure 14. The reaction catalyzed by tyrosyl protein sulfotransferase (TPST)

Figure 15. The chemical structures of rapamycin and FK506

Figure 16. Labelling reaction with N-hydroxysuccinimide (NHS) esters

Figure 17. Scheme for SNAP-tag approach

Figure 18. Structure of TGN38 and GFP-TGN

Figure 19. Topology and expression of GalT and GalT-GFP

Figure 20. Topology and expression of GalT-GFP-SNAP-Flag-His

Figure 21. Topology and expression of GalT-GFP-SNAP-Flag-His

Figure 22. BG-NHS couling to STxB

Figure 23. STxB-BG validation assay

Figure 24. Functional STxB-BG assay under conditions of inhibition of retrograde transport

Figure 25. Antibody uptake assay

Figure 26. Cell permeability test

Figure 27. Cell surface modification

Figuer 28. Verification of GalT-GFP-SNAP expression in stable cell line

Figure 29. Scheme of sulfation approach

Figure 30. Four different bSuPeRs

Figure 31. Estimation of the efficiency for cell surface modification by bSuPeR-3b

Figure 32. Cell surface modification in different pH

Figure 33. IF analysis for bSuPeR-ester and bSuPeR-R-ester pre-coupled STxB.

Figure 34. Sulfation assay of bSuPeR-ester and bSuPeR-R-ester cell surface modification

Figure 35. STxB modification by FK506-NHS in different concentrations.

Figure 36. In vitro binding between FKBP12 and STxB-FK506

Figure 37. Verification of ST-FKBP12-GFP expression

Figure 38. Capturing assay in Golgi apparatus.

TABLE OF TABLES

Table 1. Cargo proteins of retrograde route

Table 2. Proteins involved in endosome-to-Golgi transport

Table 3. Tools for analysis of the retrograde route

Table 4. Human proteins known to be tyrosine sulfated

Table 5. The most common features describing the sequence surroundings of sulfatedtyrosine residues

Table 6. The family of human FKBPs

Table 7. LC-MS identification of STxB-BG endocytosis assay

Table 8. LC-MS identification of STxB in cell surface modification assay

Table 9. Protein candidates for retrograde transport from plasma membrane to Golgi apparatus

TABLE OF SCHEMES

Scheme 1. Synthesis of BG-PEG$_9$-NHS

Scheme 2. Verification of best condition for click reaction

Scheme 3. Synthesis of FK506-NHS

ABBREVATIONS

AGT	O6-alkylguanine-DNA methyltransferase
AP	adaptor protein
APP	amyloid-β precursor protein
APS	adenosine 5'-phosphosulfate
BAD	Bcl-2 antagonist of cell death
BG	benzylguanine
bSuPeR	biotinylated sulfation site peptide reagent
CCV	clathrin-coated vesicle
CGN	cis-Golgi network
CI-MPR	cation-independent mannose 6-phosphate receptor
CPP	cell-penetrating peptide
CTx	cholera toxin
DSC	disuccinimidylcarbonate
dSTORM	stochastic optical reconstruction microscopy
ECM	extracellular matrix
Env	HIV-1 envelop protein
ERAD	ER-associated protein degradation
ERC	endocytic recycling compartment
ESCRT	endosomal sorting complex required for transport
ESI	electrospray ionization
FKBP	FK506 binding protein
GalT	β1,4 galac-tosyltransferase
Gb3	globotriaosyl ceramide
GFP	green fluorescent protein
GGA	ADP ribosylation factor-binding protein
GHS-R1a	growth hormone secretagogue receptor type 1a
GM1	glycosphingolipid monosialotetrahexosylganglioside
GST	glutathione S-transferase
HPLC	high pressure liquid chromatography
IC	intermediate compartment
IF	immunofluorescence
Ii	invariant chain
Kd	dissociation constant
LC-MS	liquid-chromatography ESI-MS
MALDI	matrix-assisted laser desorption ionization
MAP2	mitogen-activated protein kinase 2
MGMT	O6-methylguanine-DNA methyltransferase
MNK	Menkes P-type ATPase
MPR	mannose 6-phosphate receptor

MT-MMP	Membrane-type matrix metalloproteinases
MVB	multivesicular body
NEK4	NIMA-related kinase
NHS	N-hydroxysuccinimide
NSC	N-succinimidyl carbamate
PAGE	polyacrylamide gel electrophoresis
PAPS	3’-phosphoadenosine 5’-phosphosulfate
PEG	polyethylene glycol
PKA	cyclic-AMP-dependent protein kinase
PP1C	protein phosphatase 1C
RAP	Rapamycin
rER	rough ER
RS	restriction site
sER	smooth ER
SILAC	Stable Isotope Labeling with Amino acids in Cell culture
SLT	Shiga-like toxin
SNARE	soluble N-ethylmaleimide-sensitive fusion factor attachment receptor
SNX	sorting nexin
SNX-PX	sorting nexin-phox homology domain
ST	sulfotransferase
STx	Shiga toxin
STxB	Shiga toxin B-subunit
SV40	simian virus 40
TEN	tubular endosomal network
tER	transitional ER
TfR	transferrin receptor
TGN	trans-Golgi network
TPST	tyrosyl protein sulfotransferase
Vps	vacuolar protein sorting

INTRODUCTION

1. Intracellular transport pathways

The cell is the basic unit of life. It was first described by Robert Hooke in 1665. There are two types of cells: eukaryotic and prokaryotic. In this thesis project, only eukaryotic cells were studied.

Eukaryotic cells have a membrane envelop, the plasma membrane to physically separate the intracellular components from the extracellular environment. The plasma membrane has a phospholipid bilayer structure with embedded membrane proteins. Due to the large size of eukaryotic cells and unlike prokaryotic cells, the rates of chemical reactions would be limited by the diffusion of small molecules if eukaryotic cells were not partitioned into smaller subcompartments which are termed organelles. Each organelle is surrounded by phospholipid bilayer membranes and contains a unique complement of proteins which enable them to carry out distinguishing characteristic cellular functions. Transfer of material between many of these organelles, in particular, secretory or endocytic compartments, occurs by means of membrane-bound transport intermediates of vesicular or tubular shapes (Burgoyne and Clague 2003). In this process, transport intermediates bud off a donor and fuse with an acceptor compartment, consequently delivering their membrane-associated or soluble luminal constituents to the target organelle. The whole process of transfer is highly regulated by different protein machineries that control the sorting of cargos, membrane deformation, budding, translocation across the cytoplasm, and membrane fusion. The dysfunction of intracellular transport can cause serious and even fatal disease like Inclusion-cell (I-cell) disease (Plante, Claveau et al. 2008) and Griscelli syndrome (Griscelli, Durandy et al. 1978). The best explored intracellular transport pathways are in Figure 1.

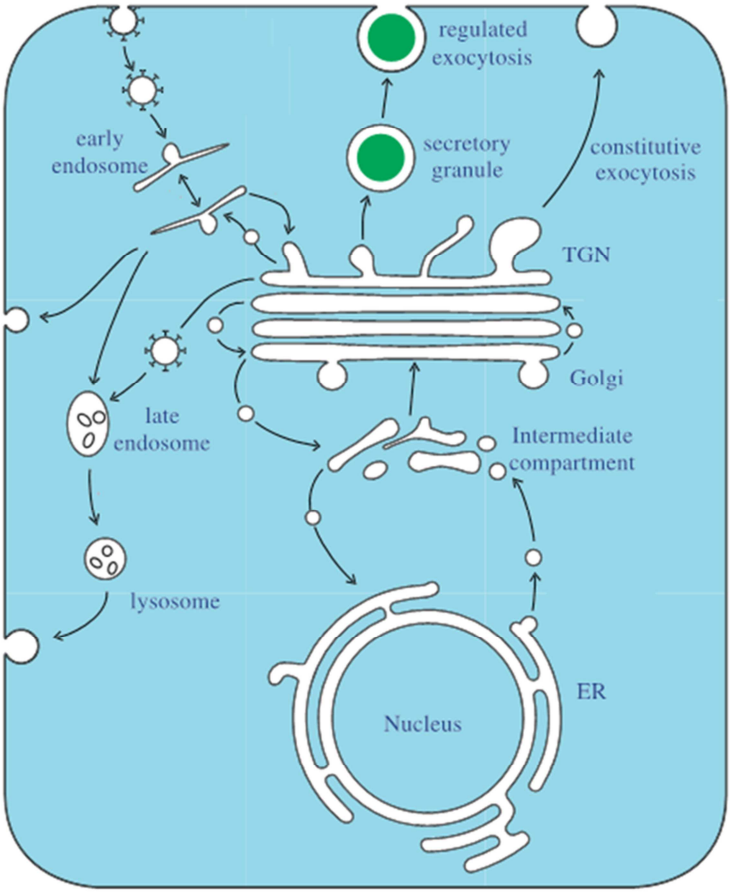


Figure 1. The different pathways of intracellular vesicle trafficking(Burgoyne and Clague 2003)

The best-studied intracellular trafficking pathways are the biosynthetic/secretory and the endocytosis pathways.

1.1 Biosynthetic/secretory pathway

The biosynthetic/secretory pathway was originally described in pancreatic exocrine cells in 1975 (Palade 1975). The pathway consists of several functionally and structurally different membrane-bound compartments which include the endoplasmic reticulum (ER), the intermediate compartment (IC), cis-Golgi network (CGN), the Golgi stacks and trans-Golgi network(TGN) as a final sorting station (Saraste and Kuismanen 1992; Harter and Wieland 1996). Eventually from TGN, matured proteins are transported to plasma membrane or other intracellular organelles. This pathway involves a series of sorting steps and a delicate balance of forward and retrieval transport regulations.

The endoplasmic reticulum is the entry station for all proteins of the biosynthetic/secretory pathway. It includes nuclear envelope, rough ER (rER), smooth ER (sER), transitional ER (tER), and intermediate compartment (IC) (Lippincott-Schwartz, Roberts et al. 2000). Secreted, membrane, and organelle resident proteins are all synthesized and processed on ER membranes. After proteins are properly processed, folded and assembled in the ER, they are further transported via the IC to the Golgi apparatus. Specific carrier vesicles bud from ER membranes and move to the cis-Golgi, via the IC (Kaiser and Ferro-Novick 1998). The complex which functions as a carrier at this step is termed COPII (Barlowe, Orci et al. 1994).

The Golgi complex, named after Camillo Golgi who first discovered it in 1898 (Golgi 1989) plays an important role at the crossroad with other trafficking pathways. Newly synthesized proteins and lipids are transported from the ER to the Golgi where they are modified and distributed to the plasma membrane or other organelles. The Golgi complex is a stack of flattened membrane sacs in which proteins and lipids are post-translationally modified during their passage. Until now, three models are used to

describe how molecules are transported within the Golgi. These are termed the vesicular transport, cisternal maturation and rapid-partitioning models (Figure 2). The vesicular transport model proposes that cargo is conveyed between successive Golgi stacks within membrane-enclosed vesicles that bud off from one cisterna and fuse with the next (Orci, Stamnes et al. 1997). In the cisternal maturation model, the Golgi apparatus is seen as a dynamic structure, in which cisternae mature and molecules are carried along passively (Bonfanti, Mironov et al. 1998). The most recent model, rapid-partitioning model proposed by Jennifer Lippincott-Schwartz indicates that proteins and lipids partition within interconnected cisternae as a consequence of their entry and exit fluxes, as well as their preference for an optimal lipid environment (Patterson, Hirschberg et al. 2008).

The trans-Golgi network (TGN) is the site of the sorting and final exit of cargo from the Golgi. It refers to the trans-side of the Golgi and structurally it is seen as a sacculotubular network (Traub and Kornfeld 1997).

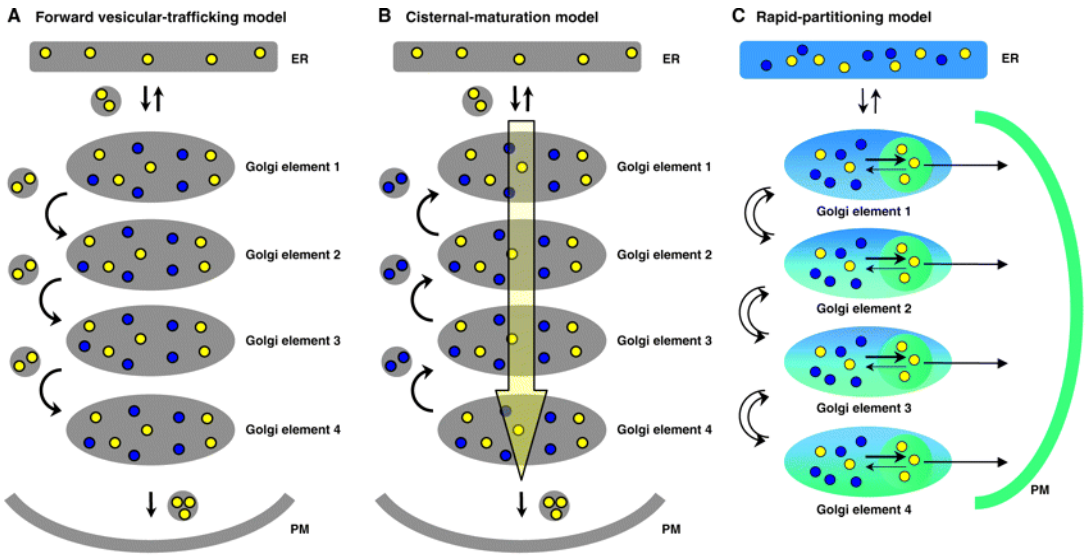


Figure 2. Models for intra-Golgi trafficking. Cargos that are synthesized in the ER and transported through the secretory pathway are shown in yellow; Golgi processing enzymes are shown in blue. Golgi element 1 is on the cis side and receives material from the ER, and Golgi element 4 is on the trans side and is involved in packaging of cargo for delivery to the plasma membrane (PM). Arrows indicate the direction of trafficking. (A) vesicular transport model, (B) maturation model, (C) rapid-partitioning model. (Bonifacino and Rojas 2006)

1.2 Endocytic pathways

As mentioned before, the cell is surrounded by a phospholipid bilayer structure termed plasma membrane which separates the cell from the outside environment. There are numerous ways that cargos and membrane proteins can be internalized from cell surface (Figure 3). The process by which cells absorb molecules such as proteins by engulfing them is defined as endocytosis. Endocytic mechanisms can be classified by their dependence of the coat-protein clathrin into clathrin-dependent and clathrin-independent.

Following endocytosis, from the plasma membrane, most internalized molecules are first transported into the dynamic tubular-vesicular network of early endosomes (Bonifacino and Rojas 2006). Internalized material is then sorted to three different destinations: late endosomes/lysosomes, the plasma membrane (recycling pathway), or the TGN (retrograde transport).

1.2.1 Late endocytic/degradation pathway

Ubiquitin and ESCRT-complexes (endosomal sorting complex required for transport) are key actors of this pathway according to the latest models (Raiborg and Stenmark 2009). Most signaling receptor proteins destined to be degraded in late endosomes/lysosomes are covalently modified with ubiquitin. The ubiquitylated cargo is captured by the ESCRT machinery to generate internal vesicles, resulting in the formation of multivesicular bodies (MVB). These MVBs migrate along the microtubules and finally fuse with late endosomes (Futter, Pearse et al. 1996). As a result of continuous acidification and constant delivery of hydrolase-containing transport vesicles from the TGN, late endosomes are eventually converted into lysosomes where the cargo proteins and receptors are degraded.

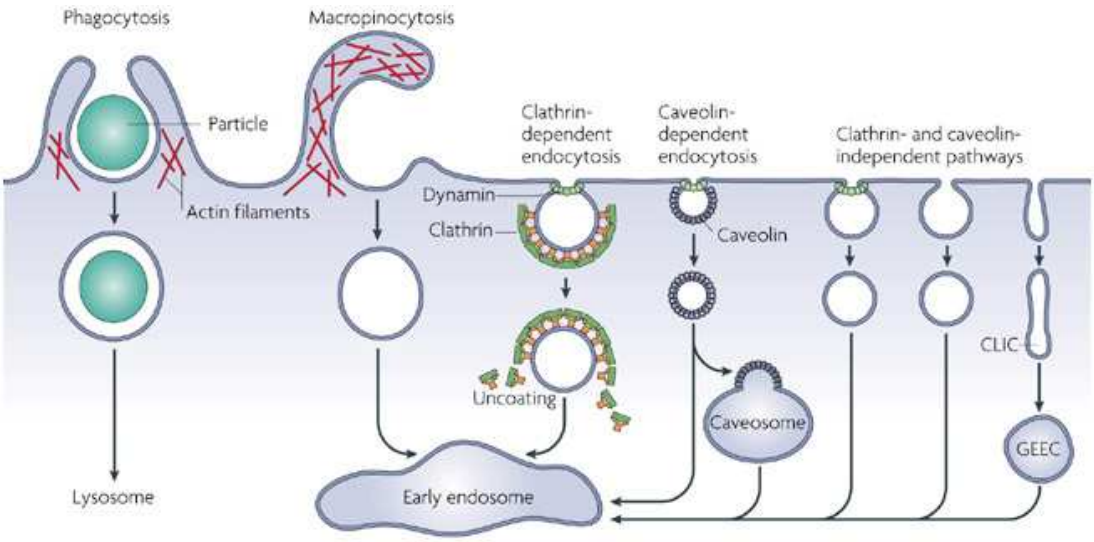


Figure 3. Endocytosis pathways into cells. Large particles can be taken up by phagocytosis, whereas fluid uptake occurs by macropinocytosis. Both processes are actin-dependent, and the size of the vesicles formed is much larger compared with the other endocytic pathways. Numerous cargoes are taken up by clathrin-mediated endocytosis, but there are a growing number of proteins that can be endocytosed by clathrin-independent mechanisms (Mayor and Pagano 2007).

The physiological importance of this pathway was first shown by studies on the extracellular signal protein epidermal growth factor (EGF) (Haigler, McKanna et al. 1979; McKanna, Haigler et al. 1979), which is taken up with its receptor into early endosomes by clathrin-mediated endocytosis and sequentially degraded in lysosomes.

1.2.2 Recycling pathway

After endocytosis, internalized proteins can also recyle back to cell surface (Figure 4). Recent work has shown that this recycling pathway is crucial in cytokinesis (Riggs, Rothwell et al. 2003), cell adhesion and morphogenesis (Paterson, Parton et al. 2003), as well as for the protein and lipid homeostasis of the plasma membrane.

Recycling can be rapid (below min) or slow (tens of minutes). Early studies indicated that the small GTP-binding protein Rab4 plays an important role in rapid recycling by controlling an early sorting events in endosomes (van der Sluijs, Hull et al. 1992). Another GTP-binding protein, Rab35, was shown to be involved in rapid recycling route (Kouranti, Sachse et al. 2006). The so-called “slow” recycling route involves the transport of cargo proteins from early endosomes to plasma membrane, via the endocytic recycling compartment (ERC) (Maxfield and McGraw 2004), which is molecularly defined by the presence of Rab11 and EHD1, and morphologically as a tubular compartment.

1.2.3 Retrograde transport

Besides the late endocytic/degradation and the recycling pathways, internalized proteins can be targeted to a third route: retrograde transport to the Golgi apparatus (Figure 5). This pathway will be presented in detail in the next chapter.

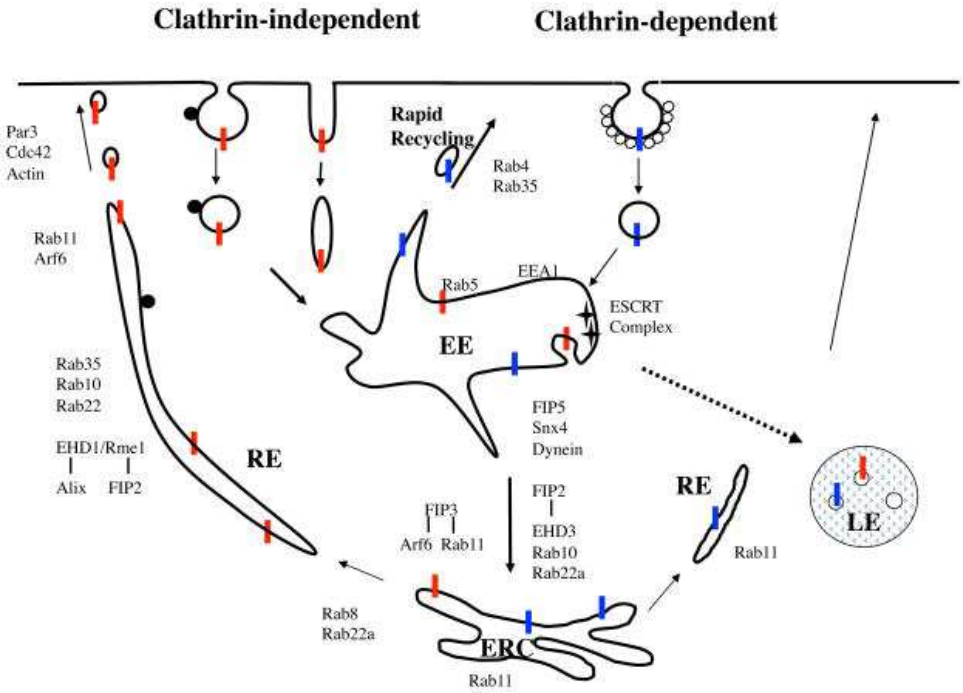


Figure 4. Recycling pathways. (Grant and Donaldson 2009).

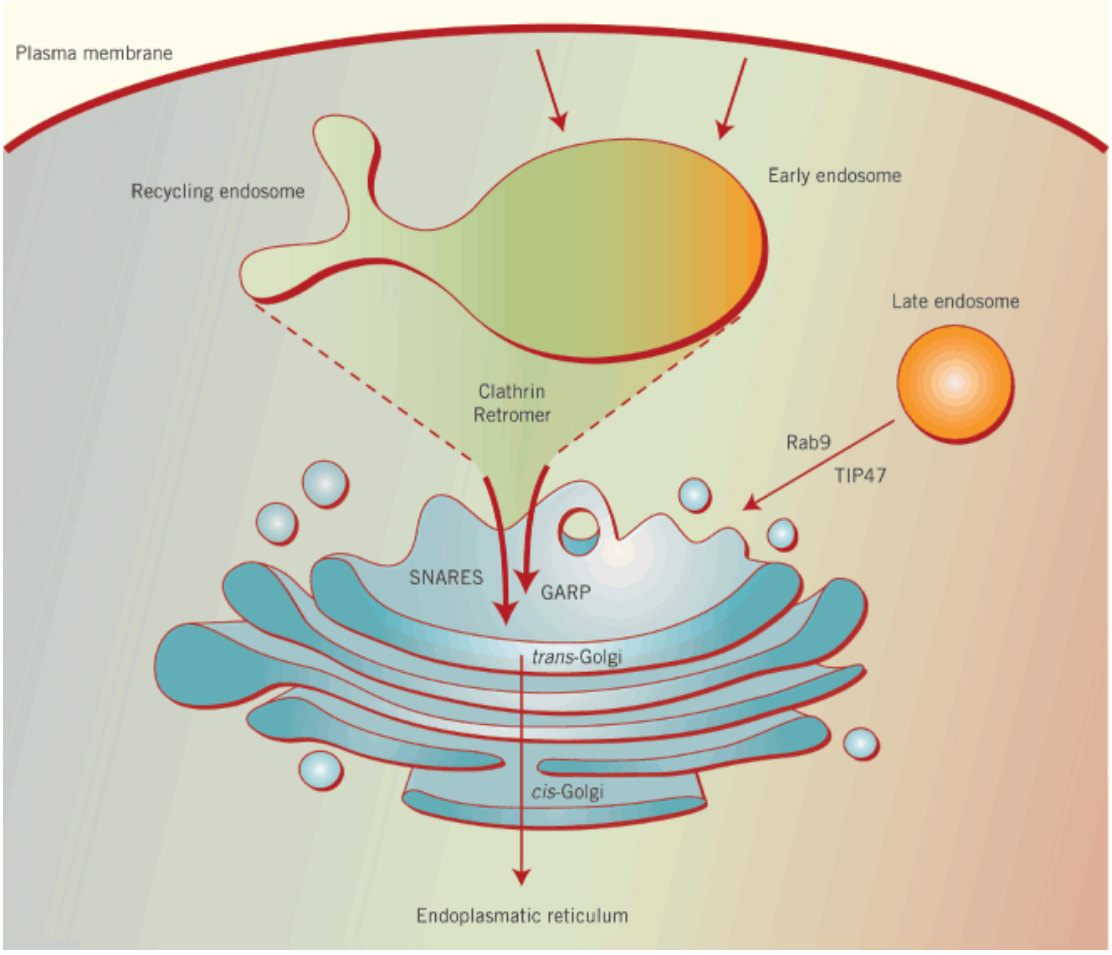


Figure 5. Scheme for retrograde transport. (Johannes and Wunder 2011)

2. Retrograde transport

The study of the cellular entry of the plant toxin ricin has led to the discovery of retrograde transport (Olsnes and Pihl 1972). An increasing number of retrograde cargo proteins could be identified in recent years, and the interest in retrograde transport has largely increased. The table shows a recent list of cargo proteins that follow this pathway (Table 1).

2.1 The functions of retrograde transport

The best-studied function of retrograde transport is the recycling of receptors which are in charge of cargo exchange between the biosynthesis/secretory and the endocytic pathways.

Mannose 6-phosphate receptors (MPRs) in mammalian cells and Vps10 in the budding yeast *Saccharomyces cerevisae* capture the newly synthesized acid hydrolases at the TGN and transport them to endosomes. Upon endosomal acidification, the ligands are dissociated from their receptors (Munier-Lehmann, Mauxion et al. 1996). The hydrolases are then transferred to lysosomes, and the receptors are recycled by retrograde transport to TGN in order to enter a new functional cycle (Duncan and Kornfeld 1988; Marcusson, Horazdovsky et al. 1994). Multiligand type-I receptors such as sortilin, sorLA and sorCS, the homologues of Vps10 in mammalian cells (Mari, Bujny et al. 2008) function similarly.

Cargo	Cellular Function	Role of Retrograde Transport in Cargo Function
Mammalian Proteins		
MPR	Lysosomal enzyme trafficking	Receptor recycling from endosomes to the TGN
Sortilin, SorLAs, SorCSs	Multiligand carriers	Receptor recycling from endosomes to the TGN
Wntless	Wingless (morphogen)-specific transport factor	Receptor recycling from endosomes to the TGN
VAMP4	SNARE	Steady-state localization
GS15	SNARE	Steady-state localization
Furin	Proenzyme maturation	Enzyme recycling from endosomes to the TGN
MT-MMP	Cleavage of extracellular matrix components	Enzyme recycling from endosomes to the TGN
P-selectin	Cell adhesion	Recycling to TGN for repackaging into new secretory granules
GLUT4	Glucose transporter	Recycling to TGN for repackaging into new storage vesicles
BACE1	Processing of beta amyloid precursor protein	Steady-state localization
BACE2	Processing of beta amyloid precursor protein	Steady-state localization
Menkes protein	Copper transporter	Functional localization to cellular compartments
Atg9	Autophagy	Functional localization to cellular compartments
EGFR	Signaling receptor	Membrane translocation
Soluble antigens	Crosspresentation	Membrane translocation
SUMF1	Posttranslational modification	Functional recycling to the ER
TGN38/46	Unknown	Unknown
Yeast Proteins		
Vps10	Vacuolar enzyme trafficking	Receptor recycling from endosomes to the TGN
Kex2	Proenzyme maturation (furin ortholog)	Enzyme recycling from endosomes to the TGN
Chs3	Chitin synthesis	Recycling to TGN, possibly for polarized secretion
Snc1	SNARE	Steady-state localization
Tlg1	SNARE	Steady-state localization
Pep12	SNARE	Steady-state localization
Fet3/Ftr1	Iron transporter	Functional localization to cellular compartments
Plant Proteins		
VSR1/ELP	Seed storage protein	Receptor recycling from storage vacuoles to TGN
Pathogen-Produced Proteins		
Shiga toxin, Cholera toxin, ricin	Toxins	Access to cytosol
HIV-1 Nef	Immune evasion	Downregulation of MHC-I, CD80, CD86
HIV-1 Env	Virus envelope	Virus assembly, surface depletion of adhesion factor
HSV-1 gM	Virus envelope	Surface depletion of adhesion factor
AAV5	Virus	Unknown

Table 1. Cargo proteins of the retrograde route. (Johannes and Popoff 2008)

The current list of proteins that cycle between endosomes and the Golgi complex includes the SNAREs (soluble N-ethylmaleimide-sensitive fusion factor attachment receptors), VAMP4 and GS15 (Tai, Lu et al. 2004; Tran, Zeng et al. 2007) which are mostly engaged in membrane fusions. Retrograde transport is also involved in development, according to recent studies. The Wnt proteins are secreted signaling molecules that play a central role in development and adult tissue homeostasis. Retrograde transport is involved in Wnt function via the trafficking of the Wnt receptor Wntless (Eaton 2008). Wntless is a conserved transmembrane protein that binds Wnt and is required for Wnt secretion. After Wntless has transported Wnt from the TGN to the plasma membrane, Wntless needs to be retrieved by retrograde transport to the TGN for a new functional cycle. Cells may also regulate retrograde transport aiming to control the range of Wnt signaling in the tissue (Yang, Lorenowicz et al. 2008).

By regulating retrograde transport, cells can balance intracellular ion homeostasis. The Menkes P-type ATPase (MNK), a copper transporter, follows retrograde transport between plasma membrane and Golgi apparatus to regulate intracellular copper levels. In cells subjected to low extracellular copper concentrations, transport of MNK from the Golgi complex is slower than retrograde retrieval from the cell membrane. MNK is consequently accumulated in the Golgi and the intracellular copper concentration is kept at normal levels. If extracellular copper concentrations are increased, the rate of MNK exocytosis is increased compared to retrograde transport so that the copper can be removed from cells (Petris and Mercer 1999). Similar function exist also in yeast for the iron transporter Fet3p/Ftr1p (Strochlic, Setty et al. 2007) whose transport is triggered by changes in iron concentration, and the mammalian glucose transporter GLUT4 which uses the hormone insulin as its trafficking signal in adipocytes and skeletal muscle cells (Shewan, van Dam et al. 2003).

Membrane-type matrix metalloproteinases (MT-MMP) are found on specialized cell surface structures, the so-called invadopodia (Poincloux, Lizarraga et al. 2009).

These proteases change their surrounding environment during angiogenesis, tissue remodeling, tumor invasion, and metastasis by cleavage of extracellular matrix components (Rowe and Weiss 2008). MT-MMP are internalized from the plasma membrane into endosomes and the TGN, and recycled back to the cell surface (Steffen, Le Dez et al. 2008). Endocytosis and recycling enables the cell to control MT-MMP activity during matrix degradation and to maintain a dynamic pool of MT-MMPs which can be used to adapt cell surface in response to physiological and pathological signals. The molecular machinery underlying retrograde transport of MT-MMP from endosomes to the TGN still needs to be identified.

2.2 Pathology

2.2.1 Shiga Toxin

Retrograde transport is used as an entry pathway for protein toxins such as Shiga toxin, ricin, and cholera toxin. Shiga toxins-producing bacteria were first described by the Japanese microbiologist Kiyoshi Shiga in 1897. Shiga toxin producing E. coli infects more than 73.000 people every year in the US alone, and has generated annual health care costs of about 400 million dollars in 2003 (Frenzen, Drake et al. 2005). The Shiga toxins family includes Shiga toxin (STx) from of Shigella dysenteriae, and Shiga-like toxins (SLT; or verotoxins) from Escherichia coli. E. coli produces different variants of two types of SLT, Shiga-like toxin 1 (Stx1; very similar to Shiga toxin) and a Shiga-like toxin 2 family of molecules (Stx2). Shiga toxin is composed of one A-subunit, consisting of the domains A1 and A2, which is non-covalently associated with one homopentameric Shiga toxin B-subunit (STxB). STxB binds the cellular toxin receptor, the glycosphingolipid globotriaosyl ceramide (Gb3) (Figure 6).

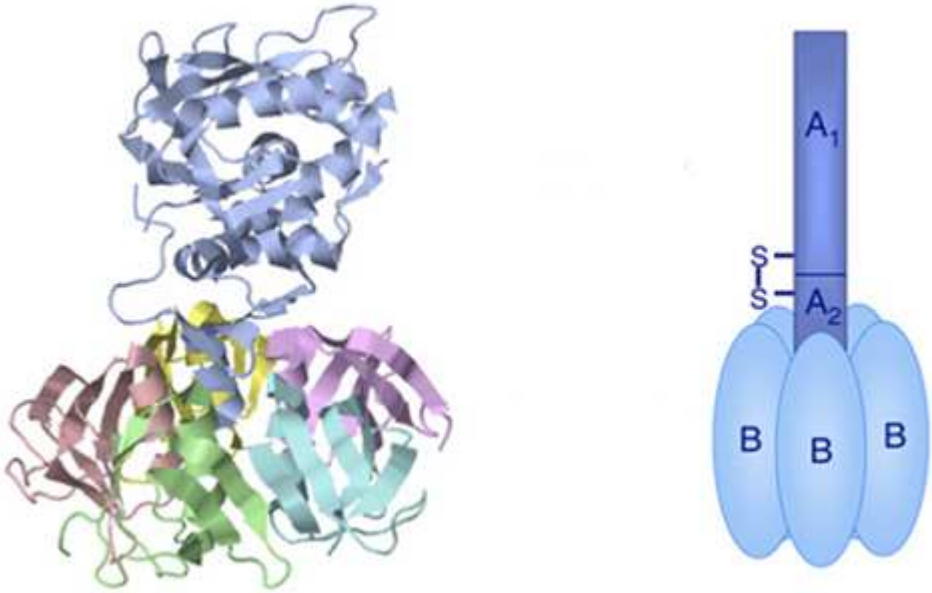


Figure 6. The structure of shiga toxin. Shiga toxin consists of an A-subunit that is non-covalently attached to a B-subunit composed of five identical fragments. The A-subunit is cleaved by the protease furin into an enzymatically active A1-fragment and a carboxyl terminal A2-fragment, which remain linked by a disulfide bond.(Sandvig, Bergan et al. 2010)

After receptor binding, Shiga toxin is internalized to early endosomes, and then transported along retrograde transport route to the Golgi, and eventually to the ER. In the ER, the cellular ER-associated protein degradation (ERAD) machinery removes misfolded, misassembled, and/or misglycosylated proteins from the ER lumen for proteasomal degradation in the cytosol (Bernasconi and Molinari 2011). Shiga toxin A-subunit uses this ERAD machinery to escape from the ER lumen to the cytosol where it cleaves a specific adenine nucleobase from 28S RNA of the 60S subunit of the ribosome, thereby inhibiting protein synthesis.

2.2.2 Ricin

Ricin is found in castor beans (Ricinus communis). Ricin is composed of the catalytic A-subunit and a monomeric B-subunit which functions for receptor binding. These two subunits are covalently linked together by a disulfide bond. Ricin B-subunit has two binding pockets (Sphyris, Lord et al. 1995) for galactose and N-acetylgalactoseamine residues of glycoproteins and glycolipids (Baenziger and Fiete 1979). Ricin is taken up into cells by multiple endocytosis pathways and transported to the ER by retrograde transport (Lord, Jolliffe et al. 2003).

2.2.3 Cholera toxin

Cholera toxin (CTx) is a protein complex secreted by the bacterium *Vibrio cholera*. This bacterium causes profuse watery diarrhea and vomiting in humans. Intestinal fluid loss is caused primarily by the release of cholera toxin, in addition to related enterotoxins, after adhesion of *Vibrio cholera* to the proximal small intestine of the host. Cholera affects 3-5 million people and causes 100,000 to 130,000 deaths a year.

The structure of Cholera toxin is similar to that of Shiga toxin: both are composed of a monomeric A-subunit and a homopentameric B-subunit. The transport route of Cholera toxin in cell is also similar to that of Shiga toxin. Instead of binding to Gb3, the B-subunit of Cholera toxin binds to the glycosphingolipid monosialotetrahexosylganglioside (GM1).

2.2.4 Virus

Besides toxins, some viruses also rely on retrograde transport. The HIV-1 envelop protein (Env) was shown to be a cargo of this pathway (Blot, Janvier et al. 2003). Depletion of retrograde transport proteins dramatically inhibits HIV replication (Lopez-Verges, Camus et al. 2006). In another study, host proteins that are required for HIV infection identified in a genome-wide RNA interference screen. This study revealed a critical role for Rab6 and its guanine nucleotide exchange factor Rgp1p in HIV-1 infection (Brass, Dykxhoorn et al. 2008). HIV immune evasion is mediated by the viral protein Nef, which induces the retrograde transport of MHC class I and costimulatory molecules (Blagoveshchenskaya, Thomas et al. 2002; Chaudhry, Das et al. 2007). The life cycle of several other viruses similarly requires retrograde transport. For example, the nonenveloped simian virus 40 (SV40) and other polyomaviruses rely on retrograde transport to the ER for their entry into the cytosolic (Spooner, Smith et al. 2006).

2.2.5 Other diseases

Some diseases like Alzheimer’s can also be related to retrograde transport. APP (amyloid-β precursor protein) is processed after endocytosis by the β-secretases BACE1 and BACE2, which cycle between endosomes and the TGN (He, Li et al. 2005). Subsequent cleavage by γ-secretase eventually leads to the formation of a 40-

or 42-residue peptide named Aβ. Excessive levels of Aβ cause serious dysfunctions and may be a reason for the progression of Alzheimer disease (Thinakaran and Koo 2008). Other studies on Alzheimer also indicate that the dysfunction of retrograde transport can be associated with late-stage-onset of the disease (Small 2008). In neuronal cells, APP interacts with the Vps10 domain-containing sorting receptor sorLA, a cycling transport of the retrograde route (see above). Genetic studies confirmed a role of sorLA in the progression of Alzheimer's disease (Rogaeva, Meng et al. 2007).

The Alzheimer-associated beta-secretase BACE1 interacts with Golgi-localized, gamma ear-containing, ADP ribosylation factor-binding proteins (GGA) that play a role in the sorting of cargo proteins from the TGN to endosomal compartments. It has been demonstrated that GGA1 also regulates the retrograde transport of internalized BACE1 from endosomal compartments to the Golgi through a direct interaction that is sensitive to phosphorylation (Wahle, Prager et al. 2005). While phosphorylated BACE1 is efficiently transported from endosomes to the Golgi, non-phosphorylated BACE1 enters the recycling route to the cell surface.

2.3 Retrograde transport machinery

Different elements of molecular machinery that is implicated in the retrograde transport as well as their functions are listed in Figure 7 and Table 2.

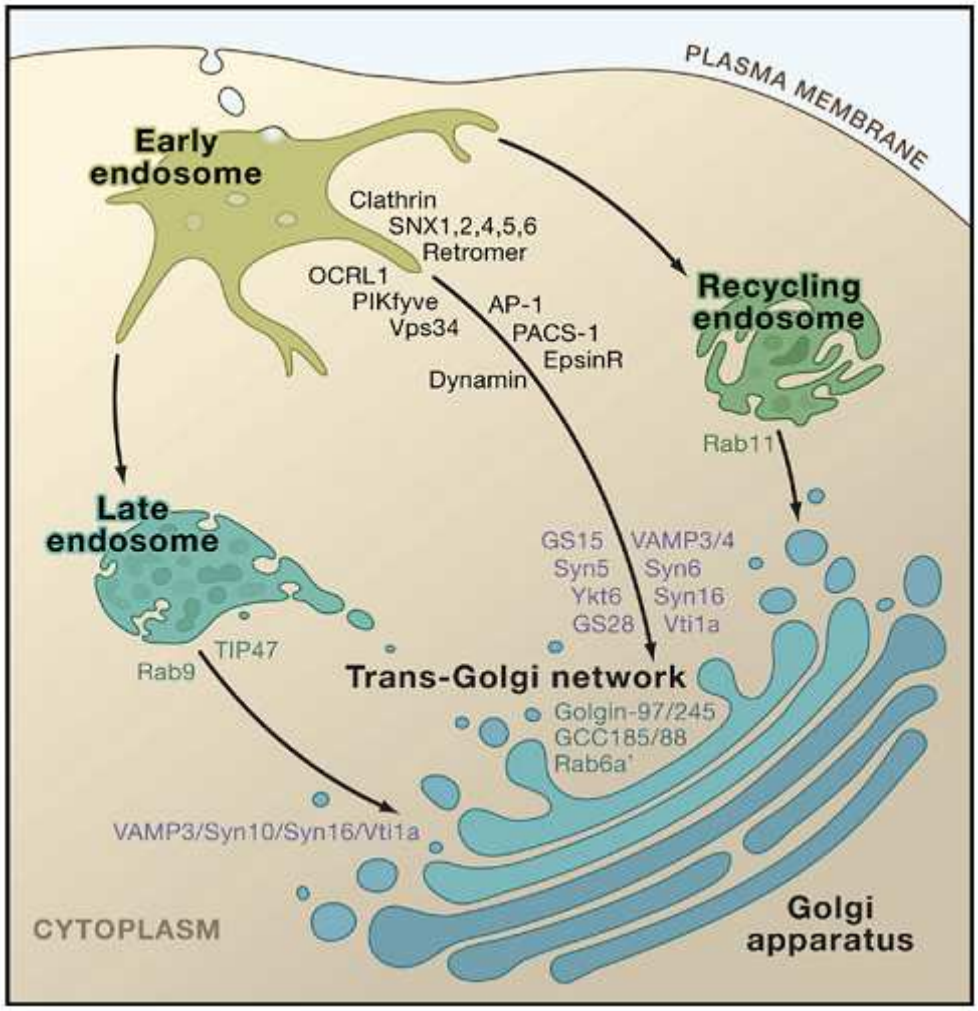


Figure 7. The location of different elements in retrograde transport (Johannes and Popoff 2008)

Table 2. Retrograde Transport at the Endosome-TGN Interface

Trafficking Factor	Localization	Proposed Function	Retrograde Cargo
Clathrin	Plasma membrane, endosomes, TGN	Nanodomain organization	STxB
AP-1	Golgi, endosomes, CCV	Adaptor	MPR, furin
AP-3	Endosomes	Adaptor	MPR
EpsinR	CCV, endosomes, TGN	Adaptor	STxB, MPR, TGN38/46, Vti1b
GGA	TGN, endosomes, CCV	Adaptor	BACE1, BACE2
PACS-1	Golgi, endosomes	Adaptor	MPR, furin, MHC-I
Dynamin	Endosomes	Scission	MPR, STxB, ricin
TIP47	Endosomes, lipid droplets	Coat	MPR, HIV-I env
Vps26/29/35 (Retromer)	Early endosomes	Coat	MPR, STxB, VSR1/ELP, Fet3-Ftr1, Wntless
SNX1	Early endosomes	Coat	MPR, STxB, sortilin
SNX2	Early endosomes	Coat	STxB, ricin
SNX4	Endosomes	Coat	Ricin
SNX5	Early endosomes	Coat	MPR
SNX6	Early endosomes	Coat	MPR
EHD1	Early endosomes	Unknown	MPR, TGN38/46
Cholesterol	Late Golgi/TGN, PM, endosomes	Membrane lipid	MPR, STxB, ricin
Golgin-97	Golgi	Tethering	STxB, CTxB
Golgin-245	Golgi	Tethering	STxB
GCC185	TGN	Tethering	MPR, STxB
GCC88	Golgi	Tethering	TGN38/46
COG	Golgi	Tethering	STxB
GARP	Golgi	Tethering	STxB, MPR
TRAPP-II	Golgi	Tethering	Snc1, Chs3
Arl1	Golgi	Tethering factor recruitment	STxB, MPR
Syntaxin 5, Ykt6, GS15, GS28 (SNARE complex)	Golgi	Fusion	STxB, CTxB, ricin, MPR
Syntaxin 6, syntaxin 16, Vti1a, VAMP3/4 (SNARE complex)	Golgi	Fusion	STxB, CTxB, ricin, MPR, TGN38/46
Syntaxin 10, syntaxin 16, Vti1a, VAMP3 (SNARE complex)	Golgi	Fusion	MPR

Table 2. *Continued*

Trafficking Factor	Localization	Proposed Function	Retrograde Cargo
Rab6a′	Golgi	Docking and fusion	STxB
Rab6IP2	Golgi	Rab GAP	STxB
Rab9	Late endosome, Golgi	Docking and fusion	MPR
p40	Unknown	Unknown	MPR
Rab11	Early/recycling endosomes, Golgi	Docking and fusion	STxB, TGN38/46
Rab22a	Early endosomes	Docking and fusion	MPR
SMAP2	Early endosomes	Arf GAP	TGN38/46
BIG1	TGN	Arf GEF	TGN38/46, furin
BIG2	TGN, recycling endosomes	Arf GEF	TGN38/46, furin
EVI5	Cytosol	Rab GAP	STxB
RN-tre	Cell surface	Rab GAP	STxB
TBC1D10A	Cell surface	Rab GAP	STxB
TBC1D10B	Cell surface	Rab GAP	STxB
TBC1D10C	Cell surface	Rab GAP	STxB
TBS1D17	Cytosol	Rab GAP	STxB
Syk	Unknown	Kinase	STxB
PKA type II α	Golgi	Kinase	Ricin
p38	Unknown	Kinase	STxB
PKCδ	Unknown	Kinase	STxB
CK2 kinase	Unknown	Kinase	MPR, furin
OCRL1	Golgi, early endosomes, CCV	PI metabolism	STxB, MPR, TGN38/46
WIPI49	Golgi, endosomes	Unknown	MPR
PIKfyve	Early endosomes	PI metabolism	MPR
Vps34	Early endosomes	PI metabolism	Ricin
Dynein	Microtubles	Motor	STxB
Microtubules	Cytoplasm/ membranes	Cytoskeleton	STxB
TMF	Golgi	Unknown	STxB

A large number of factors have been implicated in retrograde trafficking and are found at different compartments along the retrograde route. Some factors are not required for the transport of all cargos. Clathrin function is not required for MPR trafficking between late endosomes and the trans-Golgi network (TGN). STxB sorting to retrograde transport intermediates is independent of AP-1. TGN38/46 does not require PACS-1 as it does not possess an acidic cluster PACS-1 interaction motif. TGN38/46 requires GCC88 and not GCC185, whereas STxB requires GCC185 but not GCC88. Syntaxin 10 regulates retrograde trafficking from late endosomes and is therefore not needed for cargos such as TGN38/46, STxB, and CTxB, which exclusively use the early endosome-TGN interface. Ricin does not require Rab9 as it can probably use multiple trafficking receptors and pathways between endosomes and the TGN, thus in part bypassing the late endosomal pathway. CCV, clathrin-coated vesicles; CTxB, Cholera toxin B-subunit; MPR, mannose 6-phosphate receptor; STxB, Shiga toxin B-subunit.

Table 2. Proteins involved in endosome-to-Golgi transport (Johannes and Popoff 2008)

2.3.1 Clathrin coats

Clathrin is a protein that plays a major role in the formation of coated vesicles. Clathrin was first isolated and named by Barbara Pearse in 1975 (Pearse 1976). Each clathrin subunit is composed of three clathrin heavy chains and three light chains that together form a three-legged structure called triskelion (Ungewickell and Branton 1981). Clathrin triskelia self-assemble into a polyhedral cage (Blank and Brodsky 1987) (Figure 8). Clathrin can be found on the plasma membrane, the TGN, and early endosomes. It plays an important role in endocytic uptake of plasma membrane receptors. However, clathrin is unable to bind directly to membranes and needs adaptor proteins (AP) such as AP-1 and AP-2 that simultaneously bind to clathrin lipids, and protein components in the membrane (Ahle and Ungewickell 1989).

EpsinR is a clathrin adaptor that mediates retrograde transport from endosomes to the TGN. Like any member of the epsin family, epsinR comprises a phosphoinositide-binding ENTH domain. Its ENTH-domain can bind efficiently to PtdIns4P, a phosphoinositide that is found on Golgi membranes and possibly on endosomes (Hirst, Motley et al. 2003). In addition, epsinR binds to AP-1. Due to these interactions with AP-1 and PtdIns4P, epsinR is localised to endosomes and the TGN.

Clathrin coats are critical for retrograde transport from endosomes to the TGN. RNAi-mediated depletion of clathrin or epsinR leads to a block of STxB in early endosomes (Saint-Pol, Yelamos et al. 2004). Similarly, retrograde transport of MPR and TGN46 also relies on epsinR. Immunogold labeling studies revealed the presence of MPR on AP-1-containing clathrin-coated vesicles and its function in MPR trafficking to the TGN (Klumperman, Hille et al. 1993).

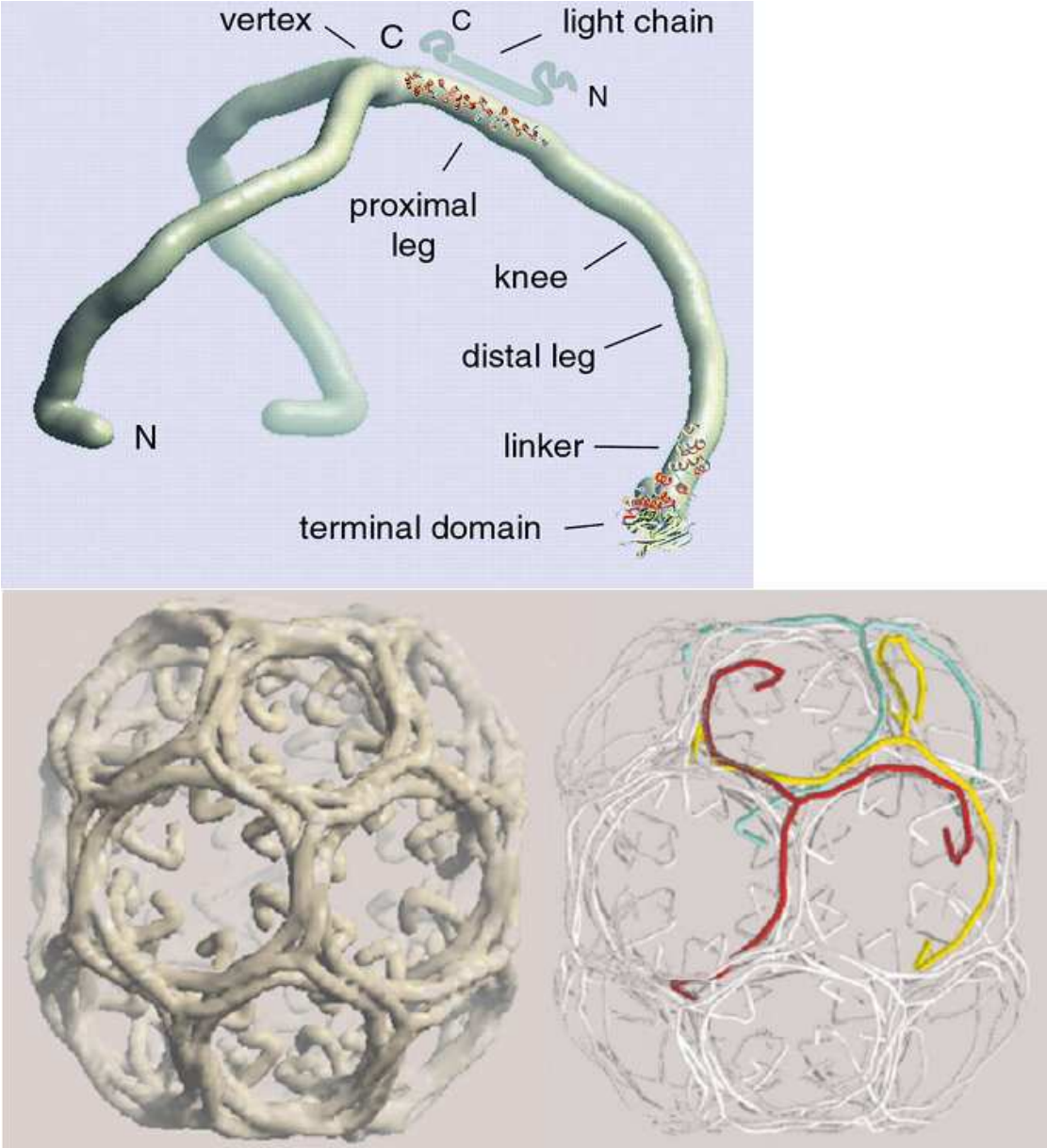


Figure 8. The structure of clathrin and the position of clathrin triskelions in a clathrin coat (Kirchhausen 2000)

2.3.2 Retromer

Besides clathrin, retromer has also been described to mediate vesicle transport between early endosomes and the TGN (Seaman, McCaffery et al. 1998). Retromer is an evolutionarily conserved protein complex composed of a trimeric vps subunit, composed of Vps26, Vps29 and Vps35, and a dimer of the BAR-domain-containing sortin nexins SNX1, SNX2, SNX5, and SNX6. Double knock-outs of SNX1/SNX2 in mice and deletion of Vps35 in Drosophila melanogaster will result in embryonic lethality (Collins 2008). These studies demonstrate the importance of retromer for cell function. SNX proteins contain a SNX-phox homology (SNX-PX) domain, by which SNX are recruited onto PtdIns3P and PtdIns(3,5)P2 enriched endosomes (Teasdale, Loci et al. 2001). The combination of phosphoinositide-binding PX and BAR-domains leads the SNX proteins to be recruited to endosomal membranes and to facilitate the formation of endosomal membrane tubules.

Retromer has been demonstrated to be critical for the retrograde transport of several cargo proteins like Wntless, MPR, and Shiga toxin (Bonifacino and Hurley 2008). For several of these cargos, clathrin and clathrin-interaction proteins are also required. A recent model explains how clathrin and retromer may function in successive steps in the formation of retrograde transport intermediates (Popoff, Mardones et al. 2007). In short, clathrin would drive membrane curvature changes and the formation of retrograde tubules. Retromer would be critical for the subsequent processing of retrograde tubules. Proteins have been identified bind both, clathrin and retromer. They could explain the molecular details of the articulation between the two machineries.

2.4 Tools to study retrograde transport

The different analysis methods to study retrograde transport are listed in Table 3. The advantage and disadvantage of those methods are also shown. The currently most popular tool in cell biology is immunofuorescence to study the intracellular distribution of toxins by microscopy, followed by an analysis of their relocalization after inhibition of retrograde transport (Johannes, Tenza et al. 1997; Amessou, Popoff et al. 2006). The toxins can be labeled with fluorescent dyes, after which they can be followed without a need for immunolabeling. Furthermore, during the initial phase of their uptake into cells, the toxins are transported in a unidirectional manner. This is different from the pool of endogenous proteins that in most cases are at equilibrium between forward and backward movements. In most cases, people use only the receptor-binding parts of the toxins in cell biology studies. These usually preserve the intracellular transport characteristics of the holotoxins without being themselves toxic.

The coupling between HRP and toxins allows the observation of the cytosolic surface of organelles by electron microscopy using a technique called the whole mount. The study of endosomal coats in the retrograde route is an example for this technique (Saint-Pol, Yelamos et al. 2004). A limitation of immunofluorescence is that it is not very quantitative. STxB-based conjugates have therefore been developed to quantitatively study the different transport steps that constitute the retrograde route. Retrograde transport to the Golgi apparatus can be measured by using a specific peptide sequence that is a recognition sequence for TGN-localized sulfotransferase. When a STxB molecule that carries a tandem of this peptide sequence, termed STxB-Sulf$_2$, reaches the TGN by retrograde transport from the plasma membrane, sulfotransferase catalyzes the transfer of inorganic sulfate from the medium onto the conserved peptide sequence. This modification can be quantitatively detected with radioactive sulfate. This technique on which one of my strategies was based will be described in detail in the next chapter.

Method	*Tool*	*Exploitation*	*Strength*	*Limitation*
Sulfation analysis	Sulfation site tagged: STxB, ricin, CTxB, MPR, antibodies	Analysis of retrograde transport to the TGN	Quantitative, sensitive, easy to use	Method does not determine absolute numbers of cargo molecule trafficking
In vitro assays	Protocols for MPR and STxB	Endosomes-to-Golgi trafficking	Immediate manipulation of cytosol activity	Difficult experimental set-up
Glycosylation analysis	Glycosylation site tagged: STxB, CTxB, ricin	Analysis of retrograde transport to the ER	Quantitative, sensitive	Kinetics of glycosylation slower than that of transport
Intoxication	STx, CTx, ricin	Analysis of retrograde transport to the cytosol	Analysis of trafficking integrating all steps of the retrograde route	Problems in calculating exact signal level due to catalytic amplification during detection
Fluorescent tagged cargos and antibody uptake experiments	STxB, CTxB and antibodies against: epitope tags (tac, CD8, GFP), TGN38/46	Visualization of compartments of the retrograde route	Dynamic analysis with morphological information	Not quantitative
Immuno-fluorescence analysis	MPR, TGN38/46, sortilin	Localization of endogenous retrograde cargo	Endogenous proteins, rapid, easily accessible	Steady-state method
Whole-mount analysis	STxB coupled to horseradish perosidase (HRP)	Electron microscopy analysis of endosomal coats in the retrograde route	High immunogold labeling efficiency	Not quantitative, no access to epitopes in organelle lumens

Table 3. Tools for analysis of the retrograde route (Johannes and Popoff 2008)

In vitro assays have been developed to exploit sulfation analysis for the reconstitution of endosomes-to-TGN transport on permeabilized cells or membrane fractions (Mallard, Tang et al. 2002). In these assays, the activity of the cytosol can be directly modified with antibodies, mutant proteins or chemical inhibitors. This powerful assay is technically difficult to handle and needs a lot of training to be mastered.

Glycosylation analysis applies a similar principle as sulfation for measuring cargo arrival in the ER using ER-localized oligosaccharyl transferase and glycosylation site-tagged STxB (Fujinaga, Wolf et al. 2003).

The sulfation and glycosylation tools can also be applied to the study of retrograde transport of endogenous cargo proteins. The different sulfation or glycosylation trafficing probes are introduced onto these cargos either directly, or via antibodies. It becomes a critical issue to know whether during the time of modification, enough endogenous cargos molecules are present at the plasma membrane to generate a signal that can subsequently be detected.

3. Proteomics

The term “proteome” was used for the first time in 1995 (James 1997) to describe the protein complement of a genome. Rapidly, the name of a new discipline was coined, proteomics. In essence, proteomics is the study of protein properties (expression levels, post-translational modifications, interactions etc.) on a large scale to obtain a global, integrated view of disease processes, cellular processes and networks at the protein level.

With the dramatically increased amounts of DNA sequences in databases, it is realized that only having complete sequences of genomes is not sufficient to elucidate biological function. A cell depends on a complicated series of metabolic and regulatory pathways for its survival. There is no strict relationship between genes

sequence and the cell functions. Proteomics is necessarily complementary to genomics because it focuses on the proteins that are the active elements in cells. Therefore, proteomics directly contributes to reveal cell functions.

Proteomics can be mainly divided into three areas: (1) protein characterization for large scale identification of proteins; (2) “differential display” proteomics for comparison of protein levels; and (3) studies of protein–protein interactions.

3.1 The functions and methods of proteomics

3.1.1 Identification and analysis of proteins

The most significant breakthrough in proteomics is the mass spectrometric identification of gel-separated proteins (Figure 9). The proteins to be analyzed are isolated from cell lysates by biochemical fractionation or affinity purification, often followed by gel electrophoresis. Usually, the proteins are then digested with a protease such as trypsin. The mass spectrometer is used to determine the masses of the digested peptides. Thereby, a mass map can be obtained. This mass map is then compared with predicted mass maps of proteins within the database to eventually identify the proteins.

The basic components of all mass spectrometers are the same. These are composed from an ionization source, a mass analyzer, and an ion detector. The ionization source is used to convert molecules into gas phase ions. For protein analysis, the two most common types of ionization sources are matrix-assisted laser desorption ionization (MALDI) and electrospray ionization (ESI). The development of ESI for the analysis of biological macromolecules (Kebarle and Verkerk 2009) was rewarded with the attribution of the Nobel Prize in Chemistry to John Bennett Fenn in 2002.

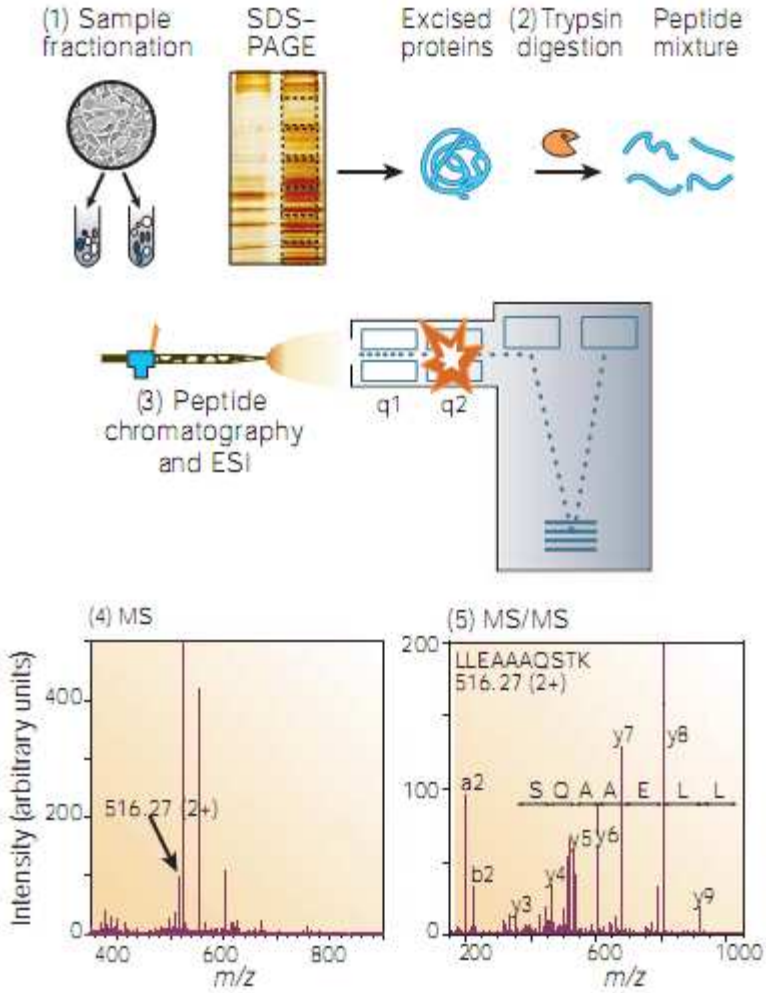


Figure 9. A strategy for mass spectrometric identification of proteins (Lueking, Horn et al. 1999)

In MALDI, the ions are created by using a laser to excite a crystalline mixture of analyte into the gas phase. ESI employs a potential difference between a capillary and the inlet of the mass spectrometer to cause charged droplets to be released from the tip of the capillary. As the droplets evaporate, gas phase charged ions are desorbed from the droplets (Aebersold and Mann 2003). ESI has advantages for constant ionization and monitoring. However, compared to MALDI, it is more sensitive to contaminating salts, buffers, and detergents. In order to fulfill ESI's stringent requirement, high-pressure liquid chromatography (HPLC) is coupled to ESI mass spectrometry constructing a liquid-chromatography ESI-MS system (LC-MS).

One of the unique advantages of proteomics studies is the ability to analyze the post-translational modifications of proteins. Phosphorylation, glycosylation and sulfation as well as many other modifications are extremely important for protein function.

These kinds of post-translational modifications can even determine a protein's activity, stability, localization and lifetime. The modification information is usually not coded by the gene sequence or mRNA expression data. Thereby, proteomics is the best choice to identify post-translational modification. However, this task is much more difficult than the simple protein identification which only requires minimal data such as one or two peptides to identify the protein in sequence databases. In post-translational modification analysis, all the peptides that do not have the expected molecular mass need to be analyzed. To fulfill the quantity demand, much more material is needed to study post-translational modifications than is required for protein identification.

Despite the disadvantage of post-translation modification proteomics, progresses have been made in recent studies, especially in the aspect of phosphorylation which is largely involved in cellular signaling pathways (Olsen, Blagoev et al. 2006). Several receptor mediated signaling pathways result in tyrosine phosphorylation of a large set of substrates. To identify these substrates, the lysates from unstimulated and stimulated cells can be prepared and resolved by 2D gels. The

proteins of interest can be detected by ^{32}P labeling or by western blotting with antibodies that recognize only the peptide with phosphotyrosine. The visualized spots can then be verified and identified by mass spectrometry since phosphopeptides are 80Da heavier than their unmodified counterparts (Soskic, Gorlach et al. 1999). An alternative strategy is to first enrich phosphoproteins by using anti-phosphotyrosine antibodies for immunoprecipitation followed by mass spectrometric identification.

3.1.2 Differential-display proteomics

Several proteomics approaches aim to discover protein functions by differential display comparison.

a). The two-dimensional gel approach

Two-dimensional (2D) gel is widely used to separate large amount of proteins. By combining 2D gel with different biological techniques like RNAi, the 2D gel can also be utilized to screen and identify new proteins and their functions. Cells are first treated in different conditions which cause distinguished proteomes. Crude protein mixture is obtained by harvesting cells. It is then applied to a “first dimension” gel strip that separates the proteins based on their isoelectric points. After the first separation, the strip is applied to a “second dimension” SDS-PAGE where proteins are denatured and separated in function of their size. The gels are then fixed and the proteins are visualized by silver staining. Unique spots in either stained gel could be of interest and are further recorded and characterized.

b). Quantitative comparison by mass spectrometry

Differential-display proteomics can also be performed using mass spectrometric

quantification. Because the intensity of a peptide peak in the mass spectrum cannot be predicted, quantification is achieved by labeling one of the two conditions by stable isotopes. Such methods have been used traditionally in mass spectrometry of small molecules, and now also been introduced into the world of proteomics.

The most recent technology in mass spectrometry quantitative comparison is named stable isotope labeling with amino acids in cell culture (SILAC). In principle, SILAC relies on metabolic incorporation of a “light” or “heavy” form of the amino acid into the proteins. Therefore, in an experiment two cell populations are grown in culture media that are identical except that one of them contains a “light” and the other a “heavy” form of a particular amino acid like ^{12}C and ^{13}C labeled L-lysine, respectively. When the labeled amino acid is supplied to cells in culture instead of the natural amino acid, it is integrated into all newly synthesized proteins. After a number of cell divisions, each particular amino acid will be replaced by its isotope labeled analog. Since there is almost no chemical difference between the labeled amino acid and the natural amino acid, the cells behave exactly the same as the control cell population grown in the presence of normal amino acid. By following the similar process as traditional mass spectrometry analysis, the “heavy” and “light” peptides will be investigated in the same spectrograph and be compared quantitatively by either Maldi-MS or LC-MS. This method can be utilized to study cell signaling, post translation modification as well as gene expression regulation.

c). Protein chips

In the protein chip approaches, a variety of probe proteins such as antibodies can be immobilized onto specially treated surfaces forming an array. Often the surface is incubated with two different groups of cell lysates which are labeled by different fluorophores and mixed so that the color itself acts as a read-out for the quantitative change of the protein binding to the antibody. This system is technically limited since it highly depends on specific and well-characterized antibodies. However, once it is

set up, protein array can provide convenient analysis. That is the reason why this method is widely used in diagnosis. In other applications, peptides, protein fragments or proteins may also be immobilized onto chips and samples applied onto the chip followed by detection of binding.

3.1.3 Protein–protein interactions

Besides individual proteins, most cell functions are carried out by protein complexes involving protein-protein interactions. Therefore, to create a protein interaction map is another goal for proteomics. Several strategies have been developed. The most popular ones are described below.

a). Purification of protein complexes

A direct way to study protein-protein interactions is to purify the entire multi-protein complex by affinity-based methods and to identify each of the components by mass spectrometry. There are multiple ways to purify protein complexes such as by using fusion proteins, antibodies, peptides, DNA, RNA, or small molecules that specifically bind to the target proteins.

Among these strategies, the common used one is to tag target proteins with an generic epitope, like the green fluorescent protein (GFP). The DNA encoding the fusion protein can then be expressed in cells, and fusion proteins can be immunoprecipitated by an antibody against the epitope together with possible interaction partners. This method requires to have access to the full-length DNA clone of the gene of interest. Because full-length cDNAs are available for all human genes, large-scale interaction studies using epitope fusion proteins have become possible.

Fusion proteins with glutathione S-transferase (GST) is another generic way to obtain interaction partners. GST is a 211 amino acid protein that catalyses the

conjugation of reduced glutathione to electrophilic centers. Instead of epitopes, the cDNA of GST is fused to the target gene. By immobilizing the GST substrate, reduced glutathione (GSH), to a solid support such as agarose, GST-tagged proteins can be pulled down, including their potential inteacting partners.

b). Yeast two-hybrid screening

Pioneered by Stanley Fields and Song in 1989, yeast two-hybrid screening was originally designed to detect protein-protein interactions using the GAL4 transcription factor of the yeast *Saccharomyces cerevisiae* (Fields and Song 1989) (Figure 10). In yeast two-hybrid screening, separate bait and prey plasmids are simultaneously introduced into yeast strain. Plasmids are engineered to produce a protein product in which the GAL4 DNA-binding domain (BD) fragment is fused onto a protein while another plasmid is engineered to produce a fusion protein in which the GAL4 activation domain (AD) fragment is fused onto another protein. The protein fused to the BD is referred as the bait protein which often is a known protein used to identify new binding partners. The protein fused to the GAL4-AD is referred to as the prey protein and can be either a single known protein or library proteins. If the bait and prey proteins interact, the GAL4-AD and GAL4-BD of the transcription factor are indirectly connected, bringing the GAL4-AD in proximity to the transcription start site and starting to transcript the reporter gene. In the contrary, if the two proteins do not interact, there is no transcription of the reporter gene. In this way, the protein interaction is linked to a change in the yeast phenotype.

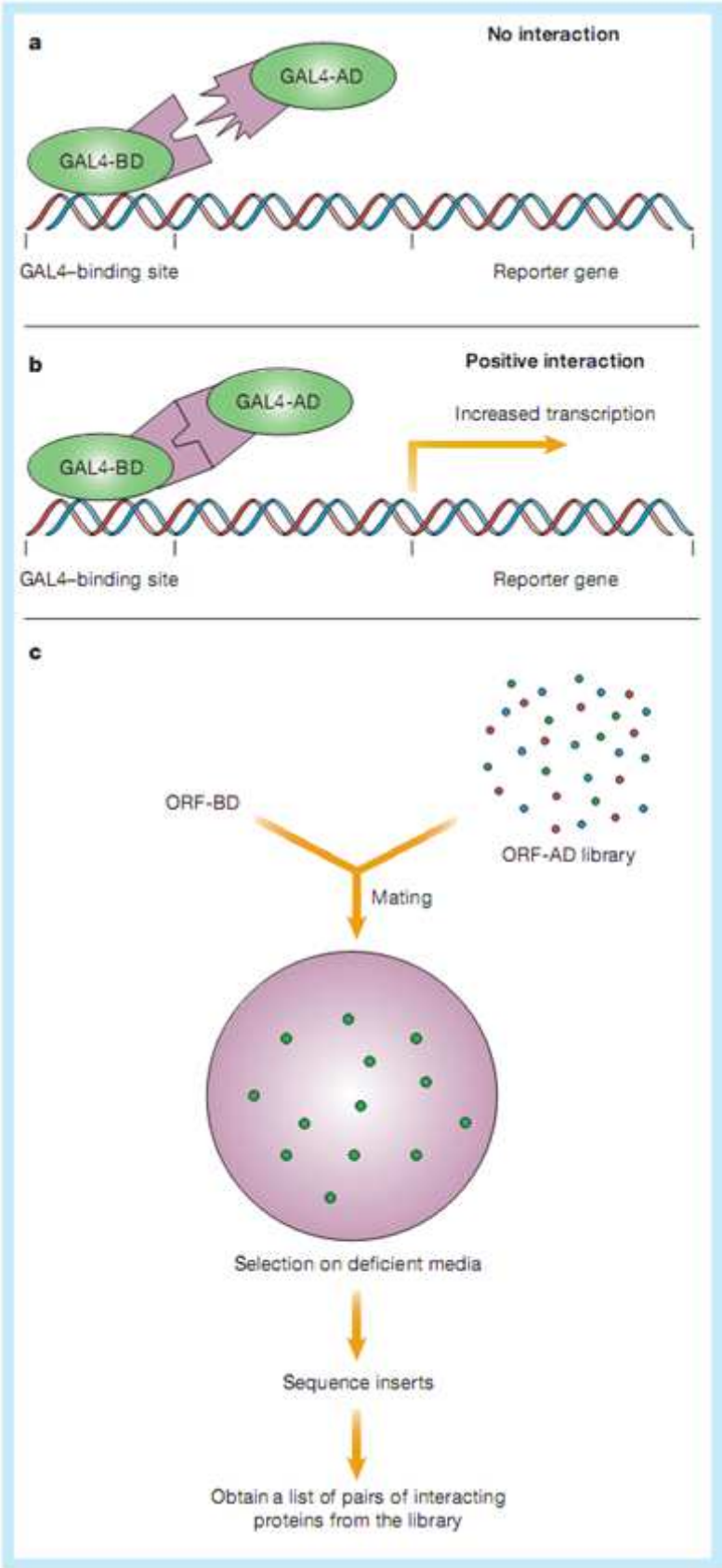


Figure 10. Scheme for yeast two-hybrid screening (Pandey and Mann 2000)

c). Phage display

Phage display is a method for the study of DNA-protein, protein-peptide, protein-protein interactions that uses the M13 filamentous phage. It was originally invented by George P. Smith in 1985. Bacteriophage particles are made to express either a peptide or protein of interest onto as a fusion with the phage's coat protein. Until now, large-scale protein interaction studies can be performed by displaying the products of cDNA libraries on phage particles. Any "bait" proteins can then be immobilized to a surface and a phage that displays a protein binding to the "bait" will remain on the surface while others are removed by washing.

Those that remained can be eluted and used to produce more phage forming a phage mixture that is enriched with relevant phage. The repeated cycling of these steps is referred to as "panning". After desired round of cycling, the relevant phage's DNA can be sequenced and the relevant protein can be therefore discovered.

3.2 Proteomics in cell biology

In order to better understand biological systems, especially at large scale, and with the goal of discovering new actors, proteomics approaches have been introduced into the realm of cell biology research. In this section, several molecular cell biology studies in which proteomics approaches were used are detailed to show how proteomics contributes to modern cell biology research.

3.2.1. A mitochondrial protein complex that links apoptosis and glycolysis

Danial N et al. provided an example of the value of mass spectrometry based proteomics for discovering connections between two independent pathways, apoptosis and glycolysis (Danial, Gramm et al. 2003). Aiming to map mitochondrial protein

complexes that contain the apoptosis regulator protein, Bcl-2 antagonist of cell death (BAD), they developed a gentle and efficient enrichment method to isolate these complexes from mouse liver mitochondria. They first isolated proteins by gel filtration according to the absorbance at 280 nm. A 2D polyacrylamide gel electrophoresis (PAGE) was followed to further separate individual proteins. A 232kDa complex which consist five major proteins was visualized by silver staining. By LC-MS, three components were identified, which were protein phosphatase 1C (PP1C), cyclic-AMP-dependent protein kinase (PKA) and the PKA-anchoring protein WAVE1. Previous study showed that PKA was an important regulator of BAD, inhibiting the activity of BAD by phosphorylating multiple serine residues (Harada, Becknell et al. 1999). PP1C functions to counter the effect of PKA, by dephosphorylating BAD (Danial, Gramm et al. 2003). To elucidate completely the function of this complex, the component WAVE1 was further characterized. Again by LC-MS analysis, the 50 kDa protein WAVE1 was identified as the glycolytic enzyme glucokinase which indicated that this complex was involved in another important cell process, glycolysis. The author proposed a model based on these results. The mitochondrial fraction of glucokinase which was associated with the BAD–PKA–PP1C–WAVE1 complex had an essential role in maintaining proper glucose metabolism function. The BAD–PKA–PP1C–WAVE1 complex functioned as a center that links cell metabolic signals and survival signals.

This early study successfully uncovered unexpected physical links between proteins that have been thought to function in independent pathways. It showed great potential in proteomics to discover novel protein complexes by analyzing and identifying protein interactions which could not be achieved by traditional cell biological strategies.

3.2.2. New component in clathrin coated vesicles

Clathrin-coated vesicles (CCV) play important roles in intracellular transport in

eukaryotic cells. As introduced above, the basic structure of CCV is composed of clathrin, which provides a scaffold, and adaptor protein (AP) complexes which select cargo and recruit clathrin to the membrane. However, the complete components of CCV are far more than clathrin and APs. Efforts were made to develop proteomics strategies aiming to identify new component in this complex (Blondeau, Ritter et al. 2004; Girard, Allaire et al. 2005). In previous proteomics studies of CCV, proteins were identified by traditional mass spectrometry methods, resulting in a long list of candidates. These studies could not distinguish which of the identified protein were true component of CCV and which were merely the contamination proteins during purification.

Borner GH et al. tried to overcome this disadvantage of protein contamination by utilizing comparative proteomics (Borner, Harbour et al. 2006). Following an established protocol (Hirst, Miller et al. 2004), CCV enriched fraction was prepared from HeLa cells. Meanwhile, another “mock CCV” fraction was purified from cells in which the clathrin heavy chain was depleted by siRNA. These cells had been shown to have no detectable clathrin coated budding profiles or vesicles (Motley, Bright et al. 2003). Thereby, the “mock CCV” should not carry any CCV components but only contamination proteins. In 2D gel, by comparing CCV and “mock CCV” spots, protein contamination could be eliminated and those unique in CCV fraction visualized. LC-MS analysis was followed to identify proteins from 63 unique spots, 28 of which had not previously been described. By analyzing these proteins, two complexes which involved in membrane trafficking were discovered, retromer and BLOC-1. This hypothesis was confirmed by western blotting against retromer and BLOC-1 subunits.

This research showed a powerful tool to investigate organelle complexes. It overcame the disadvantage of simple mass spectrometry based proteomics strategies such as protein contamination. Further, the same approach could be expanded to study the function of individual components.

3.2.3. Kinase pathways that regulate sex-specific functions in Plasmodium

Malaria is caused by parasites of the genus Plasmodium that can be transferred between vertebrates and mosquitoes by a series of life cycle (Kumar, Cha et al. 2004). A critical part in the life cycle is the generation of sexually differentiated cells, which are termed gametocytes. During the period in vertebrates, gametocytes are in stable state but when they enter mosquito, gametocytes will fertilize, producing infectious oocysts. Obviously, the proteins that distinguish male and female cells are crucial for the development of Plasmodium transmission. However these proteins had not been identified until Khan, S. M. et al. developed an innovative approach in 2005, combining cell biological technologies and proteomics methods (Khan, Franke-Fayard et al. 2005).

Before this study, it was not possible to characterize the sex-specific proteins in Plasmodium simply because of the lack of technique which can separate male and female gametocytes. Khan, S. M. et al. genetically fused the GFP sequence to a male specific or a female specific promoter, respectively the α-tubulin II or elongation factor 1α promoters. The gametocytes expressing one of these fusion proteins could thereby be distinguished and separated by FACS. The male or female cell populations were harvested, and the crude proteins mixtures were then further separated into ten fractions by one dimensional PAGE. Each fraction was digested by trypsin and analyzed by LC-MS. By comparing the male and female proteomes, new sex specific proteins were identified.

In this study, a cell biological technique to enrich distinct sexual stages of the Plasmodium life cycle was perfectly combined with mass spectrometry based proteomics. This innovation strategy leaded to the discovery of new protein kinases such as MAP2 and NEK4 that regulate male-specific and female-specific signaling pathways. This study compellingly examplifies the power of proteomics linked with cell biological methods for new pathway and protein detection.

4. Reaction schemes for proteomics approaches

During last decade, many new technologies have been developed for proteomics, as mentioned above. However, most of them are designed to identify proteins in static states such as protein complexes or cell organelles. In most cases, no dynamic information is available. In order identify the proteome of the retrograde transport route from the plasma membrane to the Golgi complex or the ER, several reaction schemes can be employed.

4.1 SNAP-tag/AGT, a DNA repair enzyme

Genomic DNA is under constant attack from exogenous and endogenous factors, especially alkylating agents. DNA alkylation even occurs spontaneously. Alkylation of DNA can occur at multiple sites. However, alkylation of the O6-position of guanine has the strongest impact. If O6-methylguanine is not repaired, it may lead to mispairing with thymine during DNA replication. In the following DNA replication cycle, thymine will pair with adenine finally causing G-C to A-T point mutations. In order to prevent this from happening, the cell is equipped with a series of DNA repair systems. Among these DNA repair enzymes, O6-alkylguanine-DNA methyltransferase (AGT) (also known also as O6-methylguanine-DNA methyltransferase, or MGMT) carries out the function to remove the methyl group at the O6-position of guanine. Strictly spoken, AGT is not a true enzyme. AGT is able to remove the alkyl group from guanine in a one-step suicide reaction (Figure 11). During this reaction, the alkyl group from O6-alkylguanine is transferred from genomic DNA onto an internal cystein residue in the active center of the AGT. This results in the recovery of guanine within the genomic DNA and leads irreversible inactivation of AGT caused by ubiquitination of the transferase. The ubiquitinated AGT is finally degraded.

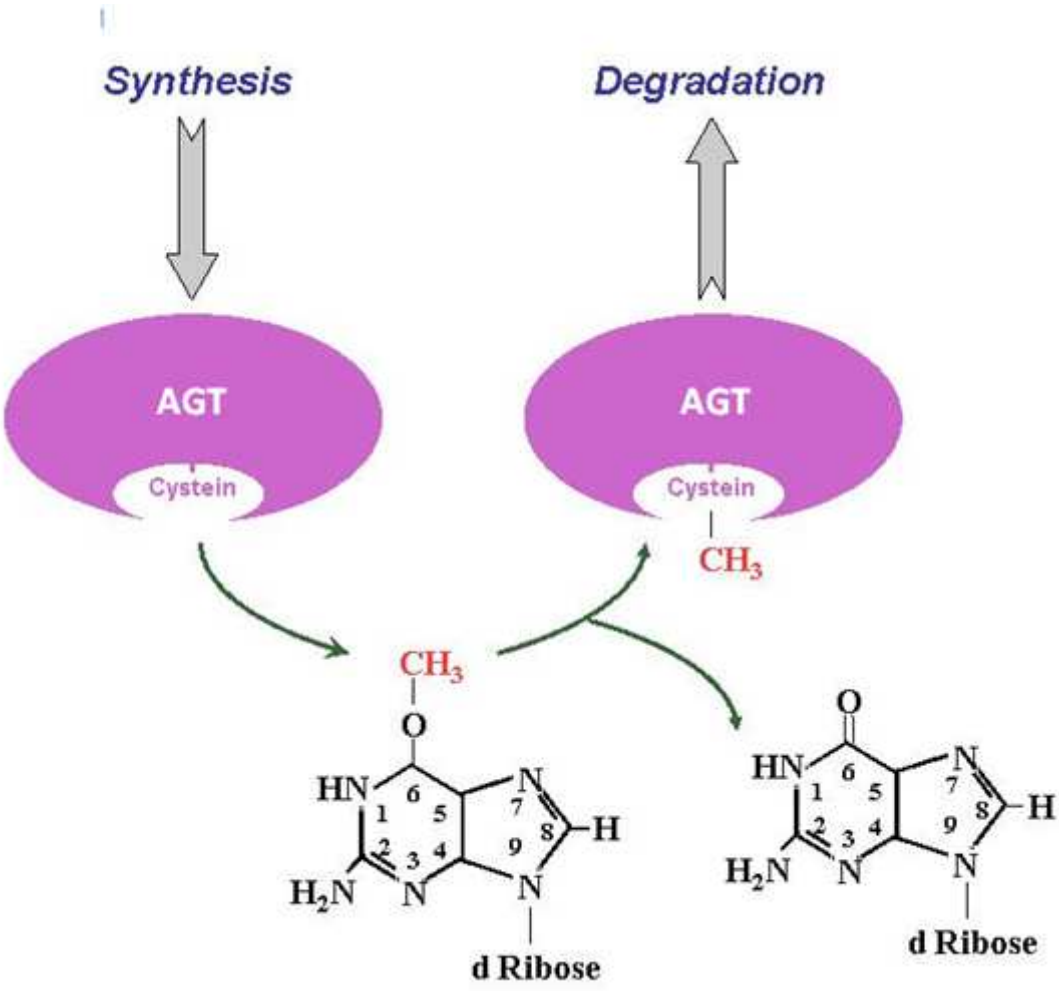


Figure 11. AGT suicide reaction

It was Johnsson K who first introduced AGT to chemical coupling in cell biology in 2001 (Damoiseaux, Keppler et al. 2001). In his study, a new substrate of AGT, termed O6-benzylguanine (BG), and several BG derivatives were synthesized and characterized. Originally, BG derivatives were synthesized with the aim of tagging endogenous AGT with biotin, exploiting the suicide reaction scheme (Figure 12). Acting like a DNA repair protein, AGT coupled itself covalently to the benzyl group releasing free guanine. In 2003, AGT was fused to other proteins and expressed in different species such as yeast, E. coli and mammalian cells (Keppler, Gendreizig et al. 2003). The AGT tagged proteins could be labeled with different chemical entities, such as fluorescent molecules, using BG derivatives (Figure 13). Johnsson K also described that the reaction rate of AGT fusion proteins with BG derivatives was independent of the nature of the label, which opened the possibility to label any given AGT fusion protein with a variety of entities. Furthermore, by using directed evolution and two different selection systems, a mutant of AGT was selected which showed an enhanced activity toward BG derivatives compared to wild-type AGT (Gronemeyer, Chidley et al. 2006). This 20 kD AGT mutant was given the name SNAP-tag.

SNAP-tag was used in various recent studies. For example, SNAP-tag was used to label subcellular structures in Zebrafish such as the nucleus, cell membranes, and endosomal system. Unlike most of the autofluorescent proteins, synthetic BG-tagged fluorophores are very stable and thereby compatible with long-term imaging (Campos, Kamiya et al. 2011). For example, a zebrafish cell lineage was tracked by fluorescence microscopy over extended periods of time. The growth hormone secretagogue receptor type 1a (GHS-R1a) could also be studied under improved conditions, using the SNAP-tag technology (Leyris, Roux et al. 2011). In combination with new microscopy methods, stochastic optical reconstruction microscopy (dSTORM), core histone H2B proteins could be visualized by fusion of SNAP-tag to H2B (Klein, Loschberger et al. 2011).

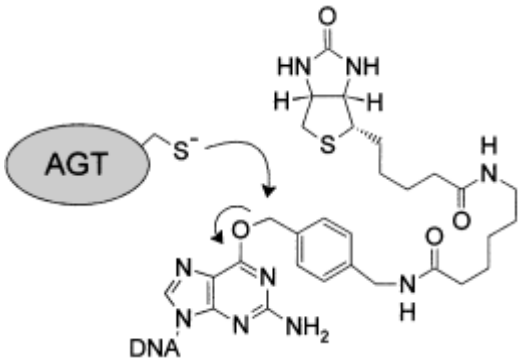


Figure 12. Coupling reaction between AGT and BG.
(Damoiseaux, Keppler et al. 2001)

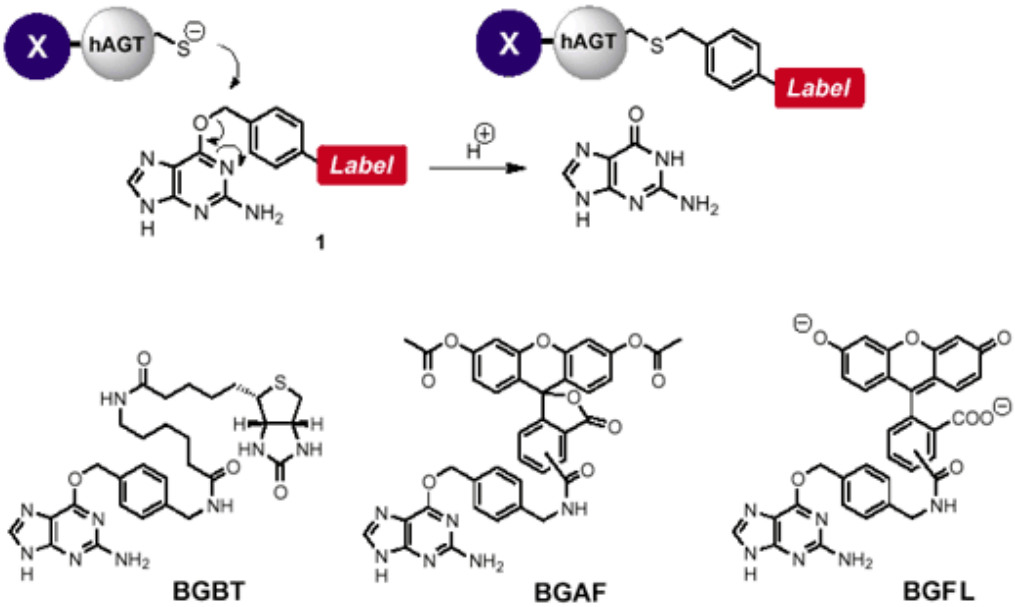


Figure 13. AGT fusion protein labeling and different BG derivatives. (Keppler, Gendreizig et al. 2003)

In principle, most of these SNAP-tag related studies employed similar strategy that SNAP-tag is fused to a protein of interest. The fusion proteins are reacted with florescent BG derivatives and can thereby be seen by fluorescence microscopy.

4.2 Sulfation

Sulfation is a chemical transfer reaction catalyzed by a family of enzymes that are termed sulfotransferases (ST). The sulfate group from the ubiquitous donor 3'-phosphoadenosine 5'-phosphosulfate (PAPS) is transferred onto a tyrosyl acceptor group of a protein recognition sequence. Mammals can produce PAPS through a pathway in which ATP and inorganic sulfate are first converted to adenosine 5'-phosphosulfate (APS) by the action of ATP sulfurylase. APS kinase then phosphorylates APS to produce PAPS (Leyh 1993). Finally, PAPS is imported into the Golgi apparatus by a specific transporter (Ozeran, Westley et al. 1996).

Sulfation plays important roles in various biological processes such as cell communication, growth and development, and defense in different species from bacteria to humans. Two groups of STs exist in mammalian cells, cytosolic STs and Golgi-membrane STs. The cytosolic STs catalyze the sulfation reaction of small molecules such as steroids, thyroids, and neurotransmitters, while the Golgi-membrane STs are involved in sulfation of large molecules including glucosaminylglycans and proteins (Negishi, Pedersen et al. 2001). In this project, we will focus our attention on protein tyrosine sulfation.

Protein tyrosine sulfation is a widespread post-translational modification of tyrosine residues (Figure 14). The sulfotransferase that catalyzes protein sulfation is termed tyrosyl protein sulfotransferase (TPST), which resides exclusively in TGN (Niehrs and Huttner 1990). Protein tyrosine sulfation was first discovered in 1954 by Bettelheim on bovine fibrinopeptide B (Bettelheim 1954).

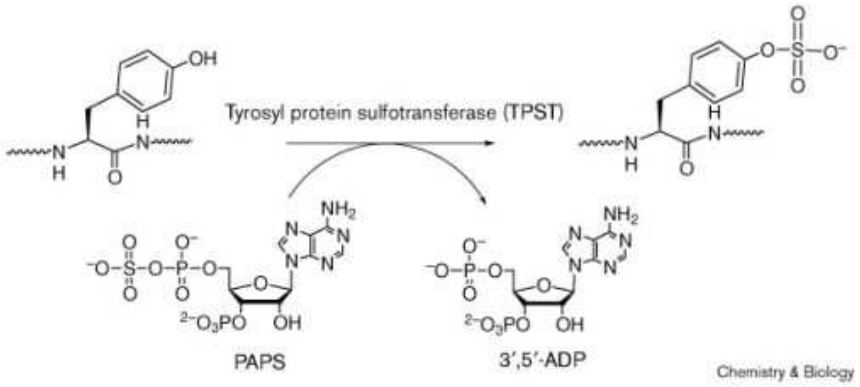


Figure 14. The reaction catalyzed by tyrosyl protein sulfotransferase (TPST) (Kehoe and Bertozzi 2000)

Like all sulfotransferases, TPST uses PAPS as a sulfate donor to transfer sulfo groups, and tyrosine is the acceptor. Significant efforts have been made to elucidate the detailed functions of the enzyme, and to predict tyrosine sulfation sites.

A number of proteins in Humans have been known to be tyrosine sulfated (Table 4). Many of these tyrosine-sulfated proteins participate in protein–protein interactions. However, other functions of tyrosyl protein sulfation still need to be further investigated.

Obviously, not every tyrosine is the substrate for TPST. Previous studies suggested that TPST substrates would not be characterized by a consensus target sequence. However, it appears that the tyrosine must be flanked by acidic amino acids to be efficiently recognized by TPSTs (Niehrs and Huttner 1990). Recent proteomics researches also confirmed this hypothesis (Table 5) (Monigatti, Hekking et al. 2006).

During early studies of tyrosine sulfation sites (Benedum, Lamouroux et al. 1987; Niehrs, Huttner et al. 1992), a specific peptide sequence Glu-Glu-Pro-Glu-Tyr-Gly-Glu (EEPEGYE) which was originally from chormogranin B, was indicated to be a substrate of TPST. An artificial gene encoding a polypeptide, termed sulfophilin, was constructed in which the sulfation sequence was repeated 12 times. This artificial gene was fused to the signal sequence of secretogranin II to direct the sulfophilin protein to the secretory pathway of 3T3 cells. A sulfation assay was developed to visualize sulfation by incubating cells with the radioactive sulfate ion $^{35}SO_4^{2-}$. The sulfated tyrosine would thus become metabolically labeled. The cell lysates were analyzed by PAGE followed by autoradiography.

Protein	Sulfation site(s), if known
Amyloid precursor	
α-2-Antiplasmin	PPMEEDYPQFGSP
CCR5	H_2N-MDYQVSSPIYDINYYTSEPCQ
Cholecystokinin	RISDRDYMGWMDF
α-Choriogonadotropin	CHCSTCYYHKS-COOH
Coagulation factor V	
Coagulation factor VIII	NEEAEDYDDDLTD
	DKNTGDYYEDSYEDS
	DQEEIDYDDTISV
	KEDFDIYDEDEN
Complement C4	MEANEDYEDYEYDELPAK
Dermatan sulfateproteogylcan	
α-Fetoprotein	
Fibrin	
Fibronectin	
Gastrin	EEEEEAYGWMDF
Glycoprotein Ib-α	EGDTDLYDYYPEEDTE
Heparin cofactor II	GEEDDDYLDLEKIFSEDDDYIDIVDS
PSGL-1	H_2N-QATEYEYLDYDFLPET
Procollagen type II	
S-protein	

Table 4. Human proteins known to be tyrosine sulfated (Kehoe and Bertozzi 2000)

Description	Residue(s)	Position(s)
Presence of one acidic amino acid within two residues of the tyrosine	Glu, Asp	− 2 to +2 Typically at − 1
Presence of at least three acidic amino acids within ±5 residues	Glu, Asp	− 5 to +5
No more than one basic residue and three hydrophobic residues	Lys, Arg, His, Leu, Ile, Phe, Trp, Met, Cys, Val	− 5 to +5
Presence of turn-inducing amino acids	Pro, Gly Asp, Ser, Asn	− 7 to +7
Absence of disulfide bonds	Cys–Cys	− 7 to +7
Absence of N-linked glycans near the tyrosine		

Table 5. The most common features describing the sequence surroundings of sulfatedtyrosine residues (Monigatti, Hekking et al. 2006)

Johannes L developed an approach that exploits this sulfation reaction to measure retrograde transport to the TGN (Amessou, Popoff et al. 2006). As mentioned before, STxB is the non-toxic subunit of Shiga toxin. STxB binds Gb3 at cell surface of target cells, and then follows the retrograde transport route to the TGN, via early endosomes. A tandem sulfation-site peptide was covalently linked to STxB (termed StxB-Sulf2), or other retrograde cargo proteins. When purified from bacteria, the sulfation site was not modified, since bacteria do not express sulfotransferase activity. The sulfation-site tagged proteins were then incubated with target cells in the presence of radioactive sulfate. Upon retrograde transport to the TGN, the sulfation-site tagged proteins would meet endogenous sulfotransferase, and their arrival in TGN membranes could then be quantified by determining the amount of radiolabeled retrograde cargo protein (STxB or others) after immunoprecipitation and autoradiography.

4.3 Rapamycin, FK506 and FKBP

Rapamycin (RAP) is a lipophilic macrolide that was identified in 1975 during antibiotic screening (Vezina, Kudelski et al. 1975). It was produced by a strain of *Streptomyces hygroscopicus*. RAP showed its immunosuppressive activity in early studies. However, it was more seriously investigated as an immunosuppressant drug since another natural product FK506 from a strain of *Streptomyces tsukabaensis* was discovered. FK506 was shown to be a powerful immunosuppressive agent as measured by its inhibitory effects in several assays of immune function (Kino, Hatanaka et al. 1987). Interestingly, structural study showed that the macrolactam rings of FK506 and RAP both contain a distinctive hemiketal-masked α, β-diketopipecolic acid amidic component (Figure 15).

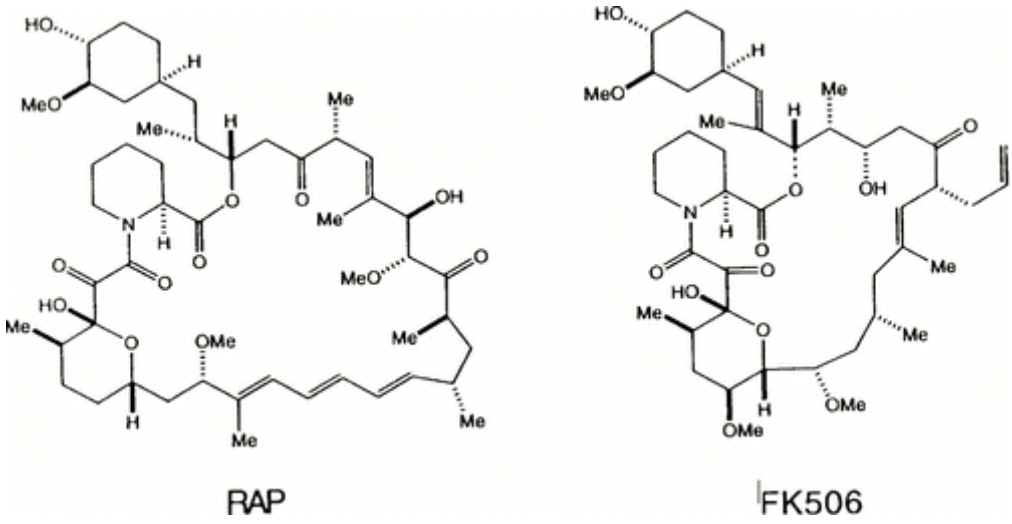


Figure 15. The chemical structures of rapamycin and FK506

RAP and FK506 bind to the same family of intracellular receptors, termed FK506 binding proteins (FKBP) (Clardy 1995; Griffith, Kim et al. 1995). Structural studies have shown that FK506 has two domains: a binding domain for FKBP, and an effector domain that, together with FKBP, forms a composite surface that interacts with calcineurin (Bierer, Mattila et al. 1990). Like FK506, RAP also has two domains: an effector domain forming a composite surface with FKBP, via a binding domain. The surface then interacts with the mammalian target of RAP, mTOR (Brown, Albers et al. 1994). The different effector domains lead RAP and FK506 to separate pathways. However, their binding domains are similar and have the same receptor, FKBP.

The human FKBP family is currently comprised of seven members (Table 6) in which FKBP12 is the most important one (Koltin, Faucette et al. 1991; Fruman, Wood et al. 1995). All FKBPs bind RAP with higher affinity than with FK506, with the exception of FKBPr38. Notably, the dissociation constants (Kd) for RAP and FK506 with FKBP12 are 0.2nM and 0.6nM respectively.

Such low Kd indicates a very strong affinity for RAP-FKBP and FK506-FKBP, which can be exploited to investigate protein-protein interactions.

4.4 Biotin and streptavidin

Biotin, also known as vitamin H or vitamin B7, is present in high amounts every living cell. In 1916, Bateman WG first discovered biotin when studying vitamins in raw eggs. Biotin's function was not fully understood until 1990s. It acts as a co-factor of carboxylating enzymes, like pyruvate carboxylase which catalyzes the formation of oxaloacetate by condensation of pyruvate and CO_2 (Diamandis and Christopoulos 1991). During the carboxylation reaction, CO_2 is first bound to the iminogroup of biotin to form carboxyl biotin, which in a second step passes CO_2 to substrates like pyruvate.

FKBP	MW	% Identity	Affinity (nM)	Affinity (nM)
	(kDa)	to FKBP12	RAP	FK-506
FKBP12	11.8	100	0.2 (kd)	0.6 (kd)
FKBP12.6[a]	11.6	83	0.2 (kd)[d]	0.5 (kd)
FKBP13	13.3	50	ND[e]	55.0 (ki)
FKBP25	25.0	40	0.9 (ki)	160 (ki)
FKBPr38[n]	38.3	33	no binding	no binding
FKBP51[b]	51.2	50	29 (IC50)[f]	166 (IC50)[f]
FKBP52[c]	51.8	53	8 (ki)	10 (ki)

Table 6. The family of human FKBPs (Abraham and Wiederrecht 1996)

Biotin is widely used for biotechnological applications, by exploiting the existence of a 53kDa protein, termed streptavidin that binds biotin with extremely high affinity. Streptavidin is a tetrameric protein purified from *Streptomyces avidinii*. The Kd of the biotin-streptavidin complex is 10^{-14} mol/L, which is one of the strongest non-covalent interactions known in nature. The affinity is about 10^{3}-10^{6} times greater than for the interaction of ligands with their antibodies. This high affinity ensures that, once formed, the complex is not dissociated by experimental manipulations such as multiple washings. Thereby, the streptavidin and biotin are widely used in protein purification, such as immunoprecipitation. Efforts were also made to express streptavidin tagged proteins in cells. However, the extremely tight binding ability to biotin makes streptavidin toxin for cells (Sano, Vajda et al. 1998). In order to express functional and less harmless streptavidin tagged protein, inducible promoters are typically used to control the intracellular amount of streptavidin (Studier and Moffatt 1986).

4.5 N-hydroxysuccinimide in protein coupling

Various methods have been developed for coupling compounds to proteins. On proteins, the preferred residue for chemical crosslinking is lysine, whose primary amino group is very reactive (Lundblad and Bradshaw 1997). The most widely used reagents are activated esters, particularly N-hydroxysuccinimide (NHS) esters (Figure 16A) (Haughland 1997).

NHS often is utilized as a cargo carrier to transfer chemical labels to amino groups in proteins (Scheurer, Roesli et al. 2005; Elia 2008). The NHS esters react at physiological pH with nucleophiles to form acylated products. The reaction can take place with several functional groups but only amines form stable conjugates (Hermanson 2008). Therefore, in proteins, mostly N-terminal and lysine residues react with NHS esters to form stable amides (Figure 16B). The reaction with amines competes with the hydrolysis of the NHS ester in aqueous solutions. The hydrolysis will be dramatically increased in basic pH. Thus, a working pH range of 7-9 is widely used to balance between NHS-amine coupling reaction and NHS hydrolysis. Besides pH, parameters such as temperature, time and the NHS ester/amine ratio need to be optimized for each reaction.

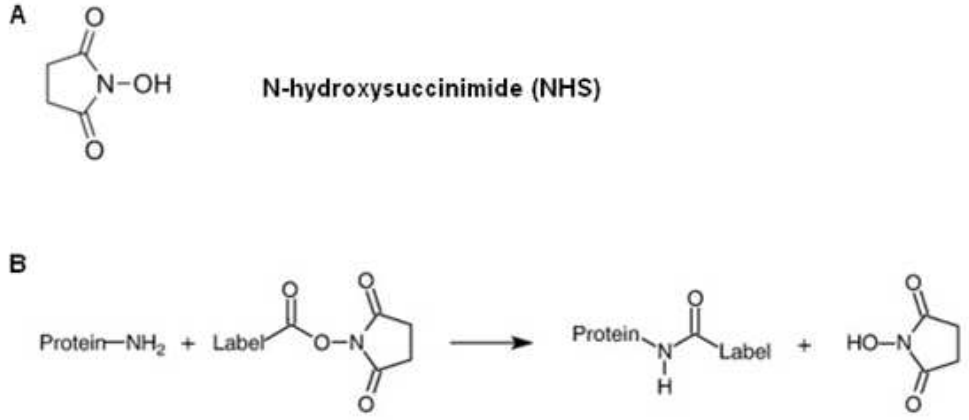


Figure 16. Labelling reaction with N-hydroxysuccinimide (NHS) esters.

RESULTS

Large efforts were made in last decade to study retrograde transport. Retrograde transport mechanisms were gradually elucidated as the list of retrograde transport factors expanded. However, the cellular functions of retrograde transport still remain largely unknown, and until now no large-scale functional proteomics studies were undertaken. In this study, we pursued four functional proteomics strategies in parallel using different types of reactions to optimize our chances of success.

1. SNAP-tag approach

Previous studies have indicated that the reaction between the SNAP-tag and BG can be applied in various contexts (Campos, Kamiya et al. 2011; Klein, Loschberger et al. 2011; Leyris, Roux et al. 2011). However, until now, the SNAP-tag was only used to visualize proteins and to trace SNAP-tag fused proteins with the help of florescent BG derivatives. Nobody had tried to achieve the reaction between full-size SNAP-tag or BG-coupled proteins in the intracellular compartments of living cells. This is what we did here to develop a novel proteomics technique of the retrograde route.

Several membrane proteins have been shown to travel via the retrograde route, such as Wntless and MPRs, similar to the well studied B-subunit of Shiga toxin (STxB) (Johannes and Popoff 2008). To identify new membrane proteins taking this pathway, a “bait and trap” strategy was adopted to the SNAP-tag technology (Figure 17). A compound termed BG-NHS was chemically synthesized: benzyl guanine, the substrate of SNAP-tag, and the N-hydroxysuccinimide (NHS) group for the coupling to plasma membrane proteins. The scheme briefly includes four steps:

a) Modification of primary amine groups of plasma membrane proteins by BG-NHS.

b) Upon incubation at 37°C, BG-tagged proteins that undergo retrograde transport are captured in the TGN by covlent interaction with TGN-localized SNAP-tag fusion protein.

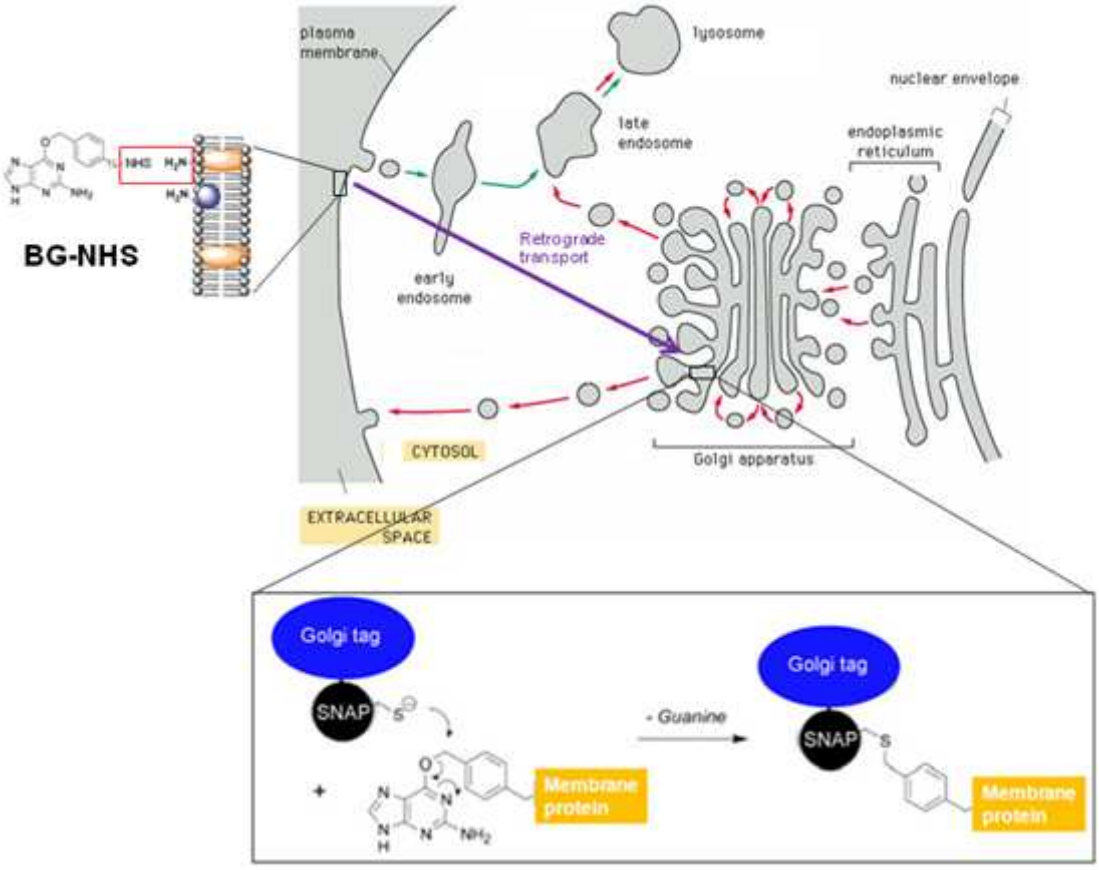


Figure 17. Scheme for SNAP-tag approach. BG-NHS couples amine group of membrane proteins in cell surface. When these BG labeled membrane proteins arrive Golgi apparatus through retrograde route, they will be captured by Golgi retained SNAP-tag. The SNAP-tag chimeric proteins are purified together with membrane proteins by immunoprecipitation and the sequences are identified by LC-MS.

c) Immunopurification of SNAP-tag protein to which BG-tagged plasma membrane proteins are covelently coupled.

d) Identification of captured plasma membrane proteins (presumed retrograde cargo proteins) by LC-MS.

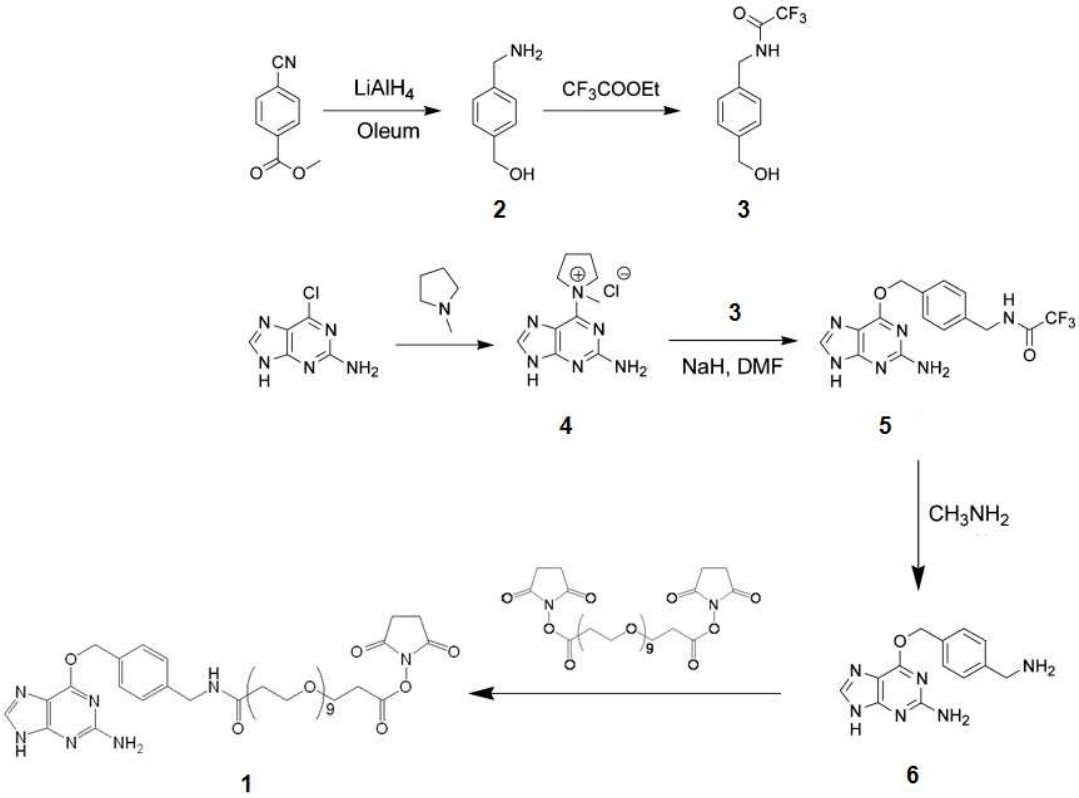
1.1 Synthesis of the BG derivative

NHS-modified BG derivatives are commercially available. These unfortunately turned out to be cell permeable (see below), and we had to synthesize such a compound ourselves. The structure that was chosen is the following: BG-PEG_9-NHS. The functional characteristics of this compound will be detailed in 1.2.2. To obtain this compound, the PEG_9-NHS group was introduced onto BG. In a first step, BG-NH_2 was synthesized according to published procedure (Keppler, Kindermann et al. 2004) (Scheme 1). BG-NH_2 was successively added to $BS(PEG)_9$, EDCI and HOBT. The speed of addition had to be controlled to avoid the formation of a BG-PEG_9 dimer. BG-PEG_9-NHS was purified in HPCL and lyophilized with a yield of 45%.

1.2 Setting-up the experimental system

1.2.1 Expressing SNAP-tag in the Golgi apparatus

In order to express the SNAP-tag in the Golgi apparatus, the cDNAs of two Golgi localized proteins, TGN38 or β-1,4-galactosyltransferase (GalT) were genetically fused the cDNA of to the SNAP-tag to obtain SNAP-TGN38 and GalT-SNAP fusion proteins.



Scheme 1. Synthesis of BG-PEG9-NHS.

a). TGN38 and SNAP-tag hybrid

TGN38 is a type I integral membrane protein that is predominantly localized to the TGN (Luzio, Brake et al. 1990). TGN38 has a single transmembrane domain. Its N terminus is in Golgi lumen and the C terminus is oriented to the cytosol (Figure 18A). The SNAP-tag cDNA was subcloned into two TGN38 plasmids provided by Banting G.

In the first plasmid, termed ΔpMEP4-GFP-TGN38, the GFP fragment is fused to the N terminus of TGN38 (Figure 18B). By sequencing we found that only one restriction site was unique and localized in the GFP sequence. This site was for PmlI. Efforts were made to insert SNAP-tag into this PmlI site. PCR was used to amplify the SNAP-tag fragment from pSNAP by primers GSL3/GSL4 bearing PmlI sites at both ends. Both SNAP-tag fragment and ΔpMEP4-GFP-TGN38 were digested by PmlI followed by a treatment of alkaline phosphatase, Calf Intestinal (CIP). However, probably due to the PmlI's blunt cutting, all the successfully ligated plasmids were found by sequencing to contain only ΔpMEP4-GFP-TGN38.

Banting G provided another TGN38 plasmid, ΔpMEP4-TGN38. A similar strategy by which GFP was inserted to TGN38 (Girotti and Banting 1996) was used to fuse SNAP-tag to the N terminus of TGN38. SNAP-tag was amplified by PCR with primer PGS7/PGS8 containing BclI sites. SNAP-tag was then inserted to ΔpMEP4-TGN38 via these BclI sites. The plasmid ΔpMEP4-SNAP-TGN38 was successfully constructed, as verified by sequencing. The ΔpMEP4-SNAP-TGN38r plasmid was then transiently transfected into HeLa cells. Cells lysate was run in SDS-PAGE and analyzed by Western blotting with antibody against SNAP-tag. It turned out that the fusion protein expression levels were extremely low (data not shown). In parallel, immunofluorescence analysis of HeLa cells transfected with the ΔpMEP4-SNAP-TGN38 and ΔpMEP4-GFP-TGN38 plasmids also confirmed the low expression of the hybrid proteins.

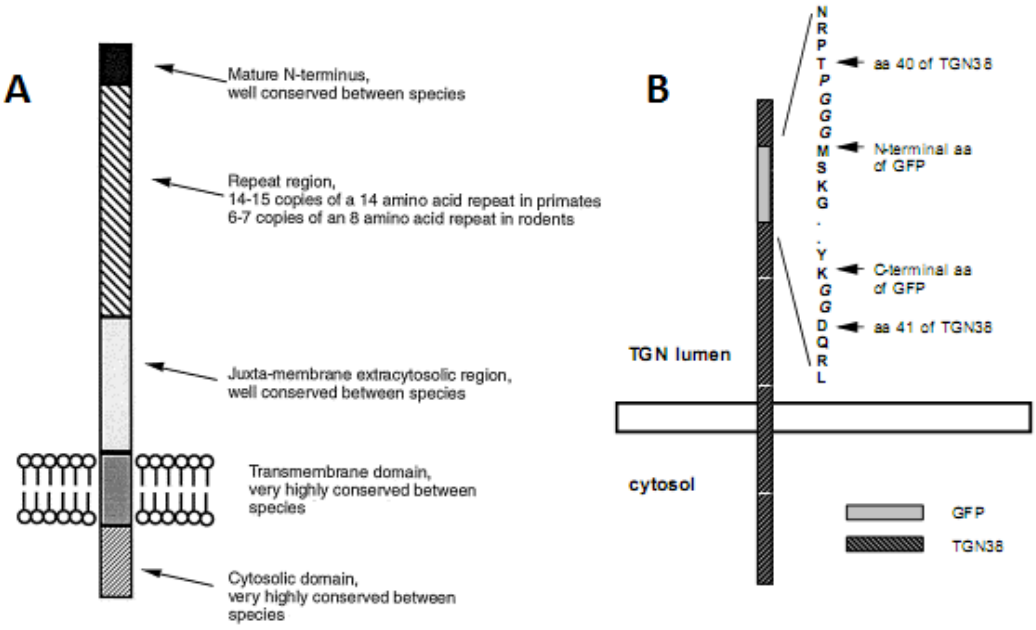


Figure 18. Structure of TGN38 and GFP-TGN. GFP is inserted to N terminus of TGN38 to ensure its TGN lumen localization.

b). SNAP-tag fusion protein with GalT

Similar to TGN38, β-1,4-galactosyltransferase (GalT) is localized in the trans-Golgi as a single transmembrane domain protein. Precisely, GalT is localized in the trans-cisternae of the Golgi (Roth and Berger 1982) and TGN (Lucocq, Berger et al. 1989; Nilsson, Pypaert et al. 1993). One notable difference between TGN38 and GalT is that the lumenal part of GalT is at the C terminus (Figure 19A). Previous studies have used a fragment of GalT to localize exogenous proteins to the Golgi apparatus (Rabouille, Hui et al. 1995). A sequence element in the N terminus of GalT (amino acid 1 to 44) was found to be necessary and sufficient to localize GalT fusion proteins accordingly. A plasmid expressing a GalT-GFP fusion protein was obtained from Franck P in which GFP was fused to the C terminus of the GalT Golgi localization sequence. In this construct, GFP was localized inside of Golgi lumen (Figure 19B). To avoid the expression problem as with TGN38, pEGFP-GalT expression was first analyzed by IF after transient expression in HeLa cells. Most cells were successfully transfected and expressed GFP (Figure 19C, green). Co-staining with a Golgi marker (GM130 in red) strongly suggested that GalT-GFP was localized in the Golgi apparatus (Figure 19C).

As mentioned in the introduction on proteomics, purification is often an essential step before protein identification. To facilitate the purification of the SNAP-tag fusion protein, two established tags (3xFlag and His) were included into its sequence. Flag is polypeptide protein tag composed of eight amino acid residues (DYKDDDDK) that is widely used in affinity chromatography to purify proteins. This short peptide has been well established and there are commercially available antibodies for both detection and purification. 3xFlag DNA fragment was amplified from commercial pCMV-3xFlag using primers PGS15/PGS17 with AscI and XhoI sites at their extremities. In PGS17, a sequence of 8 His residues was included to have a second purification handle. The 3xFlag-His fragment was inserted into the AscI and XhoI sites at the C terminus of pSNAP.

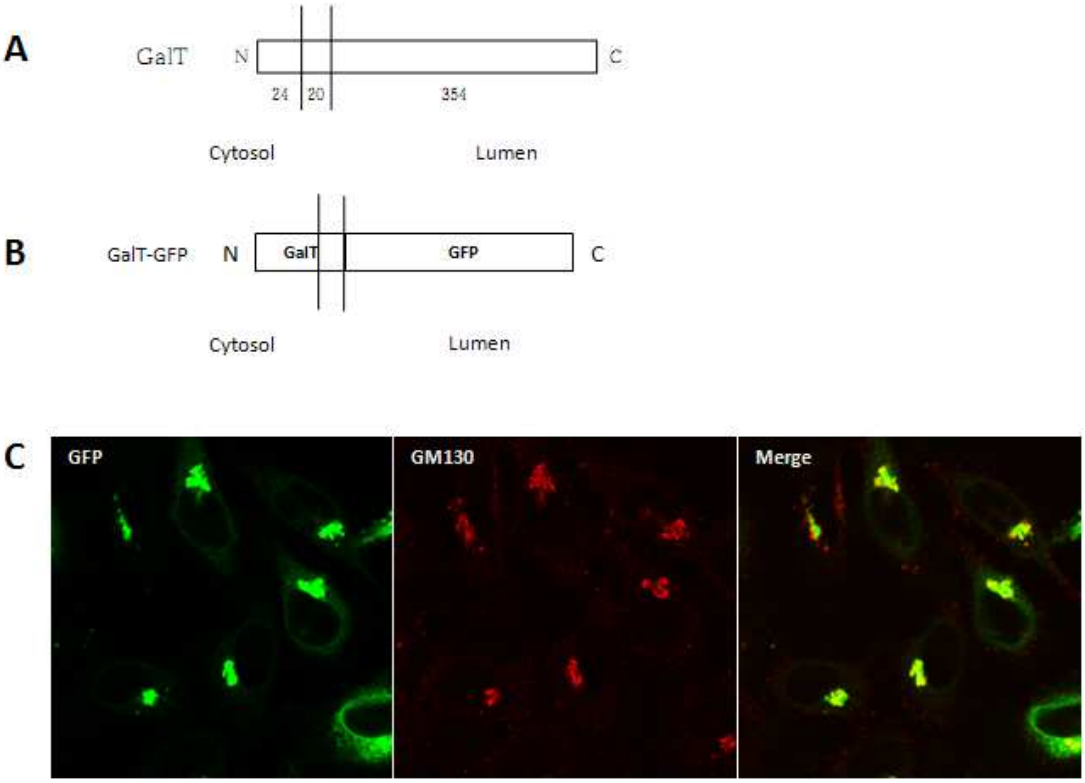


Figure 19. Topology and expression of GalT and GalT-GFP. (A)The lumenal part of GalT is at the C terminus. (B)GFP was fused to the C terminus of GalT fragment. (C)IF analysis showed the colocalization of GFP signal with Golgi marker GM130.

The entire cassette of SNAP-3xFlag-His was amplified by PGS21/PGS22 bearing BsrGI and XbaI sites, and was then inserted into the pEGFP-GalT-GFP vector at the C terminus of GFP ensuring Golgi lumen localization of SNAP-tag to the Golgi lumen (Figure 20A). The plasmid pEGFP-GalT-GFP-SNAP-3xFlag-His was transiently transfected into HeLa cells. The expected molecular weight for the GalT-GFP-SNAP-3xFlag-His fusion protein was 63.6 kDa.

Western blotting and IF analysis were used to verify the expression of the chimeric GalT-GFP-SNAP-3xFlag-His protein (Figure 20B, C). Western blotting showed a signal for GFP, but not for Flag. IF analysis also showed that GFP was correctly expressed and co-localized with Golgi marker GM130. In contrast, no specific signal could be observed with the anti-Flag antibody. It thus seemed that the anti-Flag antibody was not functioning properly.

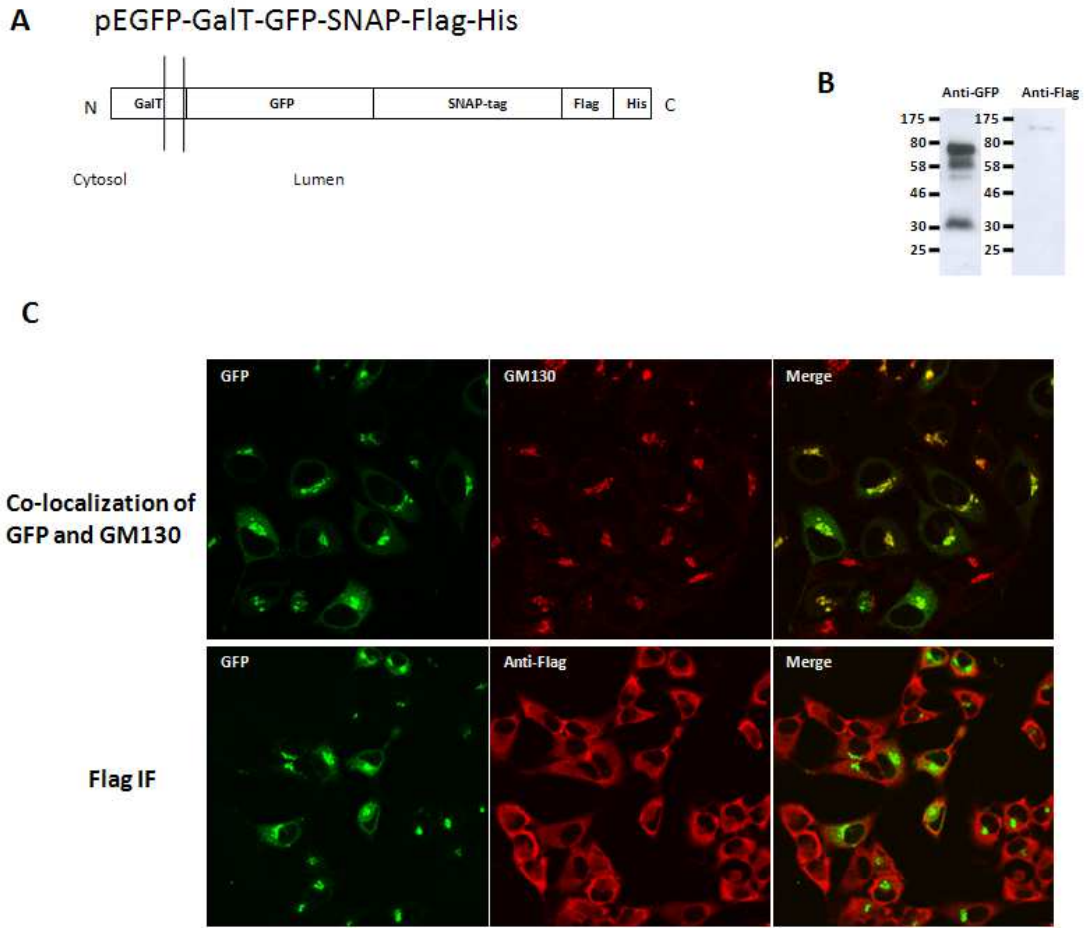


Figure 20. Topology and expression of GalT-GFP-SNAP-Flag-His. (A) The topology of fusion protein. GFP-SNAP-Flag-His was at C terminus of GalT fragment resulting Golgi lumen localization. (B) Western blotting analysis with anti-GFP antibody and anti-Flag antibody. (C) IF analysis showed colocalization of GFP signal with GM130(upper panel). No specific Flag signal was detected overlapping with GFP(lower panel).

A second fusion protein was constructed in which GalT-GFP was linked to the SNAP-tag, without inclusion of Flag or His tags. Instead of subcloning the SNAP-tag cassette onto pEGFP-GalT-GFP, the GalT-GFP cassette was subcloned into pSNAP (Figure 21A). GalT-GFP was amplified by PCR from pEGFP-GalT-GFP using primers PGS30 and PGS31 with ClaI and EcoRI sites respectively. Of note, the restriction enzyme ClaI does not digest methylated DNA, as obtained from conventional bacterial strains. To circumvent this problem, the pSNAP was amplified in a *dam*- E.coli strain. GalT-GFP was fused to the N terminus of SNAP-tag to ensure that the SNAP-tag was inside of Golgi lumen. HeLa cells were transient transfected by pGalT-GFP-SNAP vector. Cell lysate analyzed by Western blotting, which showed specific bands with antiGFP and and anti-SNAP-tag antibodies (Figure 21B). The calculated molecular weight for GalT-GFP-SNAP was 59.80 kDa, while the Western blotting revealed bands that were migrating at an apparent molecular weight that was higher than expected. A likely explanation for this migration pattern is that the fusion protein was post-translationally modified. The lower bands might be products that are obtained my proteolytic clipping. IF analysis showed that the chimeric protein co-localized with GM130, strongly suggesting that it was localized in the Golgi apparatus (Figure 21C, upper panel).

The functionality of the SNAP-tag within the fusion protein was tested using a commercially available compound named TMR-Star (Figure 21D, red).

TMR-Star is a cell-permeable red fluorescent SNAP-tag substrate that can be used to covalently label SNAP-tag fusion proteins inside living cells. TMR-Star was incubated with pGalT-GFP-SNAP transfected HeLa cells and visualized by confocal microscopy (Figure 21C). TMR-Star staining showed precisely the same pattern as the GFP, which validated both the functionality and the correct localization of the SNAP-tag fusion protein.

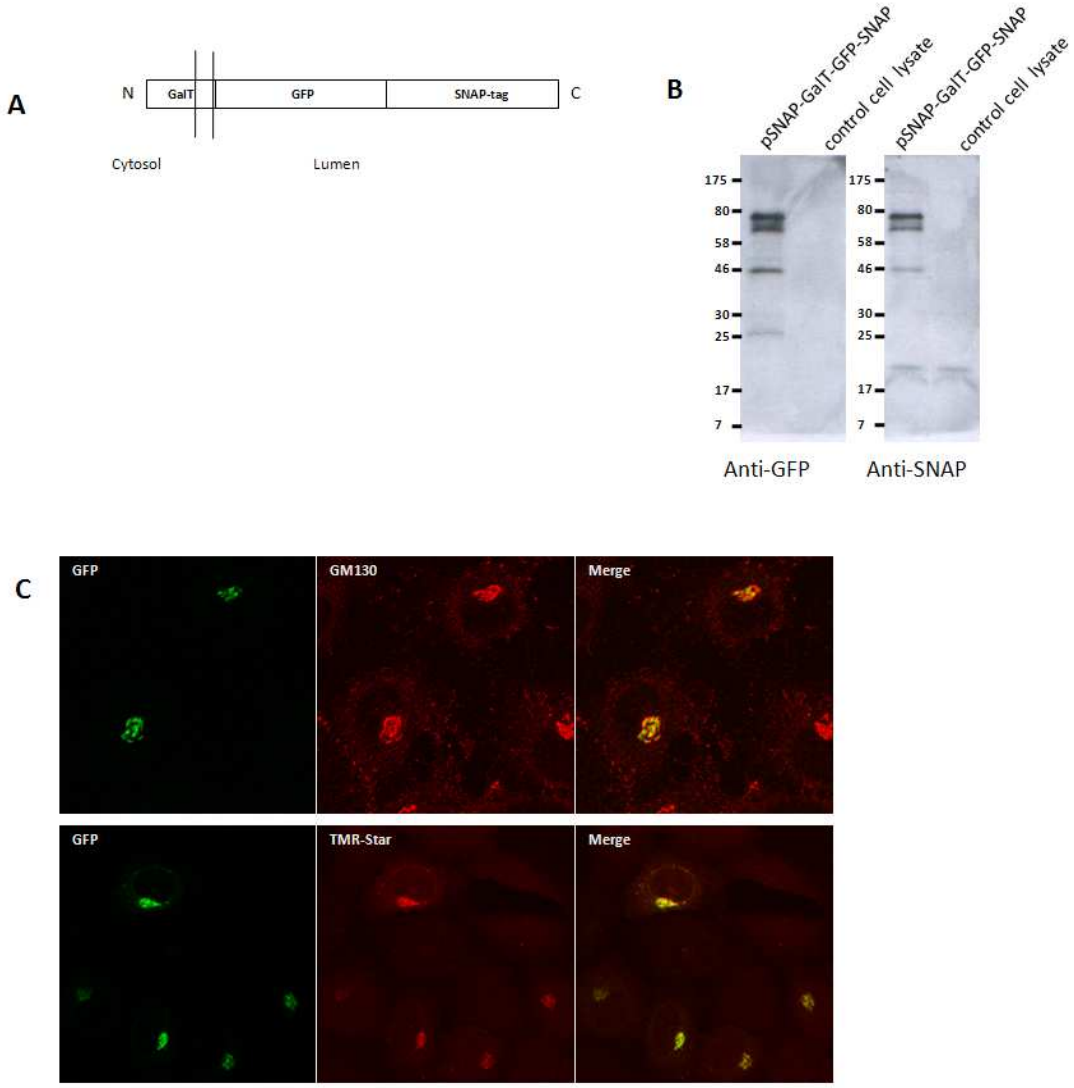


Figure 21. Topology and expression of GalT-GFP-SNAP-Flag-His. (A) The topology of fusion protein. (B) Western blotting analysis with anti-GFP or anti-SNAP antibodies. Control cell lysate was from untransfected HeLa cells. (C) IF analysis. GFP colocalized with the Golgi marker GM130(upper panel), and TMR-Star stained functional SNAP-tag which showed the same pattern as GFP(lower panel).

1.2.2 Functional validation

A series of assays were used to further assess the functionality of SNAP-tag to capture a BG-modified model protein, before venturing into large-scale proteomics.

a). STxB-BG endocytosis assay

STxB is a widely used model protein for the study of retrograde transport (Johannes, Tenza et al. 1997; Amessou, Popoff et al. 2006). The cellular STxB receptor is the glycosphingolipid Gb3. The apparent affinity of STxB for cells is very high, with Kd in the order of 19 nM (Falguieres, Mallard et al. 2001). Once it has bound to the plasma membrane, STxB can be considered as a membrane protein. STxB appeared to be an ideal model to verify our SNAP-tag approach.

A commercially available BG-NHS (MW: 481.5 Da) was coupled to STxB in PBS solution. The coupling reaction was performed in three different STxB:BG-NHS molar ratios, 1:1, 1:3 and 1:9, in RT for 2 hours. The final concentration of STxB was 150 µM. The coupling product was identified by Maldi-TOF (Figure 22). Unmodified STxB was detected as a peak at 7792 Da. During the reaction, the NHS group is released, which leads to a 367 Da increase in size on STxB. Indeed, single modified STxB at 8159 Da was observed. With increasing molar excess of BG-NHS, peaks corresponding to di or tri-modified STxB could be detected at 8526 or 8893 Da respectively.

At the ratio of 1:9, STxB was detectable at 9636.2 Da, corresponding to a modification by 5 BG groups. Each B-fragment has 5 lysines and one N terminal primary amine group. We know that over-substitution leads to a loss of STxB activity. We therefore decided to work with the coupling products that were obtained at a ratio of 1:3, in which the mono-substituted STxB was the most abundant species.

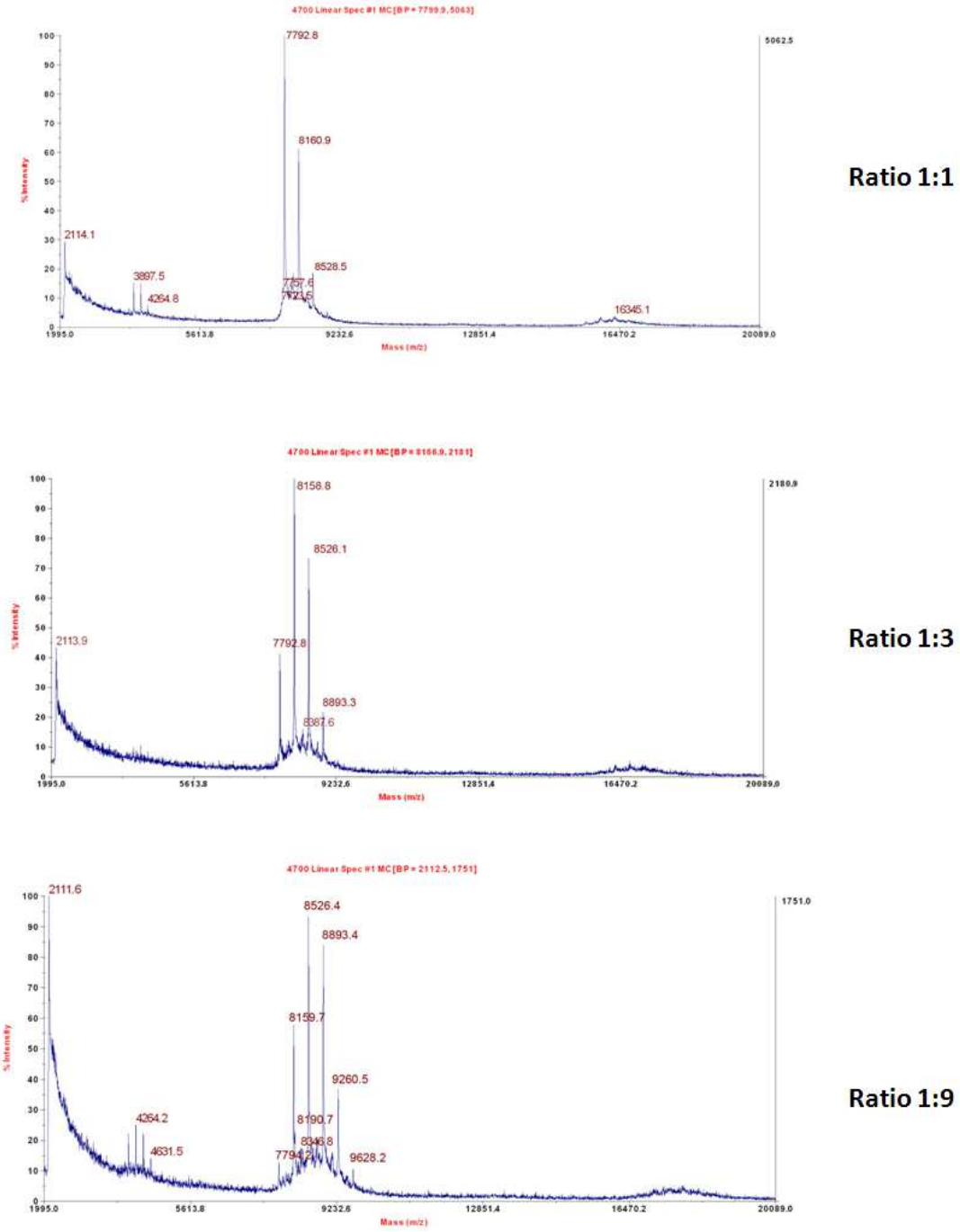


Figure 22. BG-NHS couling to STxB. STxB was coupled by BG-NHS in three different molar ratios(1:1, 1:3 and 1:9), in RT for 2 hours. Modifications were verified by Maldi-TOF/MS.

To perform STxB-BG functional assay, 400,000 HeLa cells were transiently transfected with pGalT-GFP-SNAP and incubated with STxB-BG (1µM) for 3 hours at 37°C. As controls, pGalT-GFP-SNAP transfected cells were incubated with unmodified STxB, or untransfected cells with STxB-BG. Before harvesting cells in lysis buffer, a blocking agent was incubated with cells in all conditions. This blocking agent is a cell-permeable BG derivative that quenches all SNAP-tag that has not reacted with STxB-BG, thereby preventing capture of STxB-BG by pGalT-GFP-SNAP after lysis. As described before, the SNAP-tag coupling reaction with BG is irreversible. Therefore, the blocking agent can inactivate SNAP-tag without reversing interactions between STxB-BG and SNAP-tag that preexisted to lysis. The protein GalT-GFP-SNAP (coupled or not to STxB-BG) was purified from cell lysates by immunoprecipitation using anti-GFP antibodies. After SDS-PAGE, Western blotting membanes were probed with antibodies against STxB. If STxB-BG was coupled to SNAP-tag, the band representing the GalT-GFP-SNAP-STxB protein should be at 67.6kDa. The Western blotting revealed a band at about the expected size (Figure 23; note that the band is slightly higher, likely due to glycosylation of GalT). In contrast, this band was not observed in the control conditions.

To confirm that the GalT-GFP-SNAP-STxB protein was formed after retrograde transport of STxB-BG to the Golgi, pGalT-GFP-SNAP transfected cells were incubated with STxB-BG at 37°C in the presence of the small molecule retrograde transport inhibitor Retro-2 (Stechmann, Bai et al. 2010), or at 4°C.

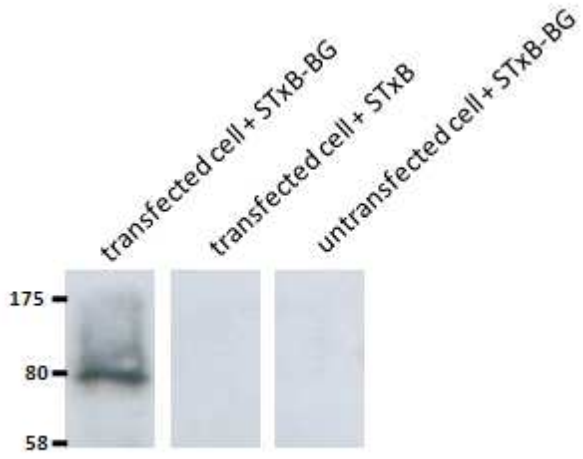


Figure 23. STxB-BG validation assay. STxB was coupled to BG-NHS. The STxB-BG(1μM) was incubated with pGalT-GFP-SNAP transfected HeLa cells for 3 hours at 37°C. After quenching, the cells were lysed and GalT-GFP-SNAP protein species were immunoprecipitated with anti-GFP antibody. Western blotting was carried out using anti-STxB antibody. The other two lanes were control conditions without either GalT-GFP-AGT transfection or BG modification.

Retro-2 selectively blocks retrograde toxin trafficking at the early endosome-TGN interface. pGalT-GFP-SNAP transfected HeLa cells were pre-treated for 30 min with Retro-2 at a final concentration of 20 μM. The compound was then kept present during all subsequent steps until cell lysis. In the other condition, pGalT-GFP-SNAP transfected cells were kept in 4°C to completely block endocytosis and intracellular trafficking. As shown in Figure 24, the band corresponding to GalT-GFP-SNAP-STxB was dramatically decreased in the Retro-2 condition (lane 2), and basically absent in the 4°C condition (lane 3). This experiment proved that the formation of the GalT-GFP-SNAP-STxB protein depends at least to a great extent on STxB-BG trafficking from the cell surface to the TGN.

b). LC-MS detection

LC-MS was used in the conditions as described above to confirm the molecular identity of the band at about 67.5 kDa, and demonstrate that our approach could lateron work in a proteomics setting to identify unknown proteins. The experiment was performed with STxB-BG and pGalT-GFP-SNAP transfected HeLa cells at 37°C, as described above. After SDS-PAGE, the area of the gel around the expected size was cut, incubated with trypsin, and peptides released from the gel were injected into the LC-MS setup. Peptides from STxB as well as from GalT-GFP-SNAP were identified (Table 7). This result thus demonstrated that our SNAP-tag approach could be used to detect and identify by LC-MS proteins that were captured via the GalT-GFP-SNAP.

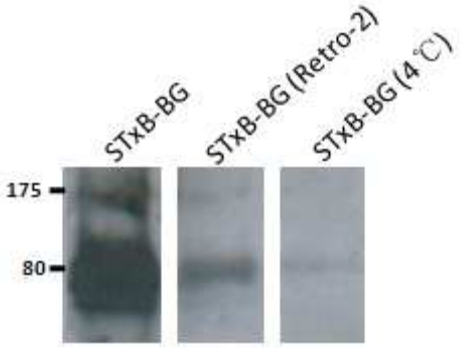


Figure 24. Functional STxB-BG assay under conditions of inhibition of retrograde transport. The STxB-BG(1μM) was incubated with pGalT-GFP-SNAP transfected HeLa cells for 3 hours at 37°C. After quenching, the cells were lysed and GalT-GFP-SNAP protein species were immunoprecipitated with anti-GFP antibody. Western blotting was carried out using anti-STxB antibody. In Retro-2 treatment condition, retrograde transport inhibitor Retro-2(20 μM) was added to culture medium to block STxB from trafficking to Golgi apparatus. In 4°C control condition, cells were kept at 4°C to completely block endocytosis.

Identifier	Size (aa)	Mol. weight (Da)	Match group	Score	Matching peptides
					All
GalTGFP	366	39608	GalTGFP	131.73	3
STxB	70	7793.8	STxB	88.48	2
SNAP	193	20239.3	AGT	101.94	1

Table 7. LC-MS identification of STxB-BG endocytosis assay. Peptides from STxB as well as from GalT-GFP-SNAP were identified.

c). Antibody uptake assay

STxB is an exogenous protein for mammalian cells. As a first step to validate the SNAP-tag approach on endogenous proteins, we used an antibody uptake protocol.

As described in the introduction, the cation-independent mannose 6-phosphate receptor (CI-MPR) shuttles newly synthesized mannose 6-phosphate tagged lysosomal enzymes from the TGN to the endosomal pathway (Ghosh, Dahms et al. 2003; Arighi, Hartnell et al. 2004). After delivering its cargo, the empty receptor returns to the TGN by retrograde transport. A small fraction of the CI-MPR is also found in the plasma membrane, where it binds to and internalizes external ligands for delivery to lysosomes (Lin, Mallet et al. 2004). To validate our SNAP-tag approach on MPR, we used a HeLa cell line that expresses a GFP-tagged version of CI-MPR (Waguri, Dewitte et al. 2003). We have previously shown that addition of an anti-GFP antibody to these cells leads to the piggyback trafficking of the antibody from the plasma membrane to the TGN. The anti-GFP antibody was conjugated to BG-NHS, as described above for STxB. The conditions were taken from a previous study in which the same antibody had been crosslinked to a sulfation peptide (Amessou, Popoff et al. 2006). pGalT-GFP-SNAP was transfected into these cells, which were then incubated with anti-GFP-BG(15 µg/ml) for 4 hour at 37°C. Extracellular antibody was removed by PBS washing. SNAP-tag was inactivated by the blocking agent, and the cells were lysed. The antibody was pulled down on Protein G Sepharose beads. The samples were analyzed by Western blotting with antibody against GFP and SNAP (Figure 25). Since the antibody modification can occur in any amine group including light or heavy chain, DTT was not added to sample buffer to observe entire antibody. A band corresponding to the size of GalT-GFP-SNAP-antibody (219.8kDa) conjugate was detected in the standard condition (lanes 1 & 5). When the anti-GFP antibody was not BG modified, no band at the size of the full size conjugate could be detected (lanes 2 & 6) while the weak band below might be contamination.

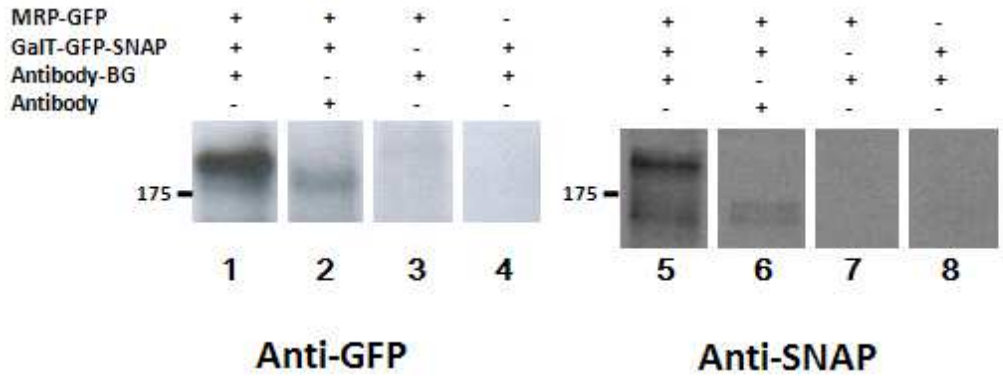


Figure 25. Antibody uptake assay. Anti-GFP-BG/anti-GFP antibody were incubated with pGalT-GFP-SNAP transfected/non-transfected HeLa/MRP-GFP cells for 4 hour at 37°C. Extracellular antibody was washes away. After quenching cell lysate was incubated with Protein G Sepharose beads to pull down antibodies together with GalT-GFP-SNAP. Western blotting was performed using both anti-GFP and anti-SNAP antibody.

As a further control, the GFP-CI-MPR cell line was not transfected with pGalT-GFP-SNAP. Again, no bands of the right size could be found under these conditons (lanes 3 & 7). Finally, we used wild-type HeLa cells that lack the SNAP-tag bait. Again, no bands were seen (lanes 4 & 8). With these control conditions, the antibody uptake assay clearly indicated that the SNAP-tag in Golgi was able to capture BG coupled antibody, most likely after its GFP-CI-MPR dependent retrograde transport.

d). Cell permeability test

As it is indicated in the scheme of Figure 17, the proteomics version of the SNAP-tag approach relies on cell surface modification by BG-NHS. For this to work, one needs to make sure that the BG-NHS is membrane impermeable. Otherwise, it would diffuse into the cell and inactivate the TGN-localized SNAP-tag before any retrograde transport can occur. The plasma membrane is negative charged and has a structure of phospholipid bilayer with a hydrophobic core, we condidered two strategies to prevent BG-NHS from passive crossing by diffusion: introducing negative charge or bulky hydrophilic groups. We encountered problems to synthesize a negatively charged version of BG-NHS due to the reactivity of the NHS group. As an alternative strategy, a polyethylene glycol (PEG) linker was added to the BG molecule. PEG is a polyether compound that is often used as a spacer arm to maintain conjugate solubility. We needed to determine how many ethylene glycol repeats are required to provide sufficient hydrophilicity such as to prevent BG diffusion across the plasma 4°C membrane. To do this, a permeability assay was developed. Cells were incubated at 4°C for 1h with BG-NHS (commercial compound), BG-PEG_3 (commercial compound), BG-PEG_9-NHS, or a blocking against that is conceived to inactivate SNAP-tag, even on live cells. At 4°C, endocytosis is completely blocked. The external compounds were then washed away, and cells were shifted to 37°C for 15 min to facilitate the reaction between SNAP-tag and these compounds if they were

cell permeable. Cells were then incubated with a cell-permeable compound named TMR-Star, a red fluorescent substrate of SNAP-tag, at 37°C for 30 min. TMR-Star conjugation to SNAP-tag could only occur if the prior incubation with BG compounds has not quenched the active center of SNAP-tag. Cells were then lysed, cell lysates were subjected to SDS-PAGE, and the gel was directly visualized using a fluorescence scanner (Figure 26). The first two lanes were controls to indicate maximum and minimum signal levels, using non transfected (lane 1) and pGalT-GFP-SNAP (lane 2) transfected HeLa cells. The addition of the blocking agent led to complete inactivation of the SNAP-tag, since no fluorescence signal could be detected (lane 3). Incubation with BG-NHS had the same effect (lane 4), indicating that cell membranes were maximally permeable to the compound. A small amount of SNAP-tag remained intact after incubation with BG-PEG3 (lane 5, arrow), and incubation with BG-PEG_9-NHS (lane 6) did not affect the TMR-Star at all. It could be concluded that the BG-PEG_9-NHS is membrane impermeable.

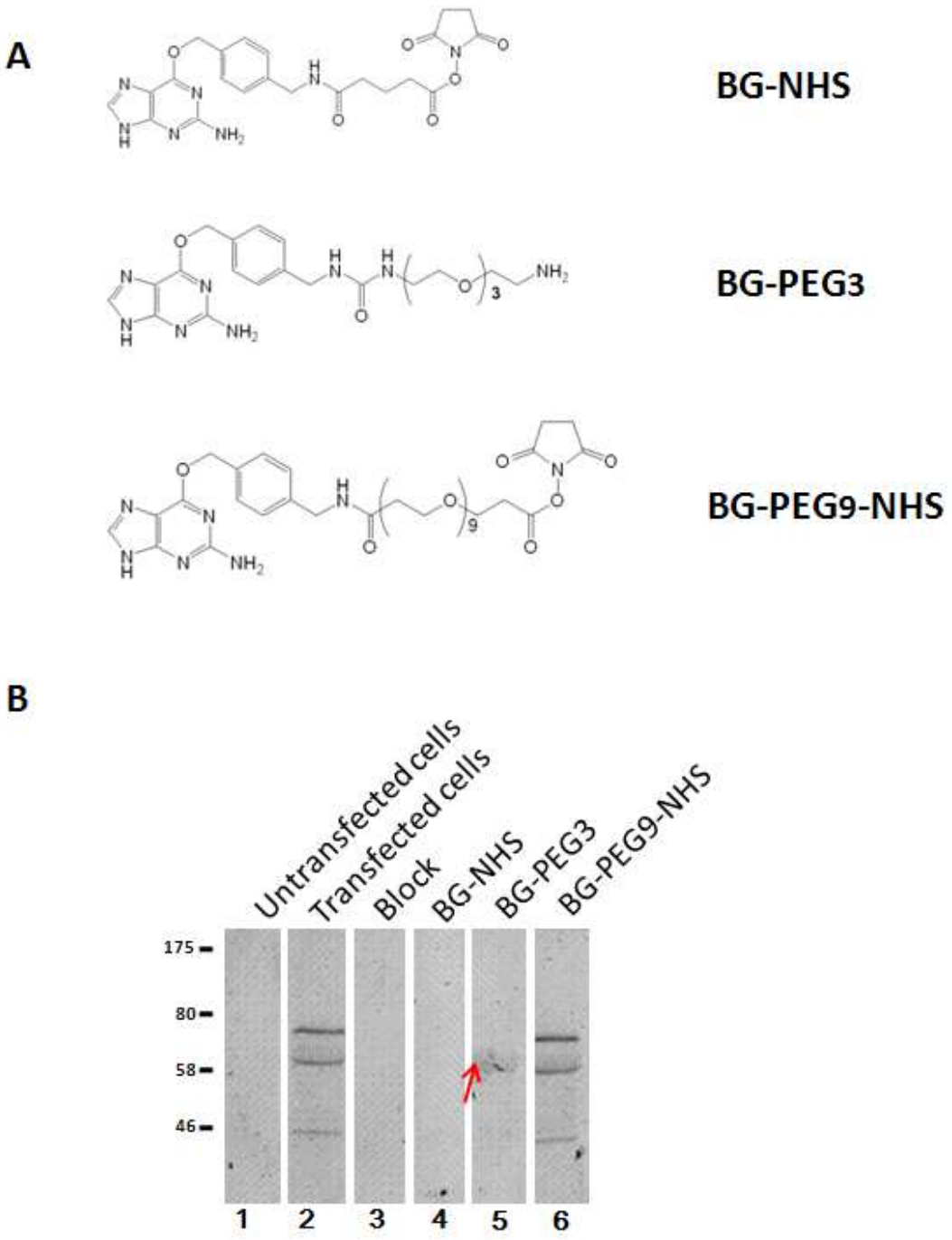


Figure 26. Cell permeability test. (A) Three BG derivatives with NHS, PEG_3 and PEG_9. (B)The three BG derivatives as well as blocking reagent were incubated with pGalT-GFP-SNAP transfected cells at 4degree for 1h. The external compounds were washed away, and cells were shifted to 37°C for 15 min. Cells were then incubated with a cell-permeable compound TMR-Star at 37°C for 30 min. Cell lysates were subjected to SDS-PAGE, and the gel was directly visualized using a fluorescence scanner. Lane 1 and 2 used untransfected and transfected cells without BG derivatives treatment to provide minimum and maximum signals.

e). Cell surface modification

The validation experiments that have been presented above used STxB or antibody conjugated with BG that was obtained *in vitro* before addition to cells. Screening for new retrograde transport cargo proteins requires cell surface modification by coupling the BG group to plasma membrane proteins. To validate feasibility, STxB was bound to pGalT-GFP-SNAP transfected cells at 4°C for 45min. At this temperature, STxB binds to its receptor, Gb3, without being internalized. These cells were incubated on ice with different concentrations of BG-PEG$_9$-NHS, and then shifted to 37°C for 3 hours. After blocking residual SNAP-tags, the fusion protein was immunoprecipitated with anti-GFP antibodies, and immunoprecipitated material was analyzed by Western blotting with antibodies against STxB (Figure 27). As a positive control, cells were incubated with BG-STxB (lane 3), and as a negative control, unmodified STxB was bound to cells that were not incubated with BG-PEG$_9$-NHS (lane 4). At both concentrations of 0.6 mM (lane 1) and 2 mM (lane 2), cells that had been surface modified with BG-PEG9-NHS showed the same band density as those had been incubated with the *in vitro* conjugate BG-STxB. This finding suggested that basically all cell surface bound STxB had been modified by BG-PEG9-NHS, indicating that the procedure was very efficient.

LC-MS analysis was again performed to confirm the protein sequence identity in the band that was visualized by Western blotting as the GalT-GFP-SNAP-STxB conjugation product. Peptides from GalT-GFP, SNAP, and STxB could be detected (Table 8). The identification by LC-MS of STxB after cell surface modification suggested that our SNAP-tag approach was ready for functional proteomics to find new cargo proteins of the retrograde route.

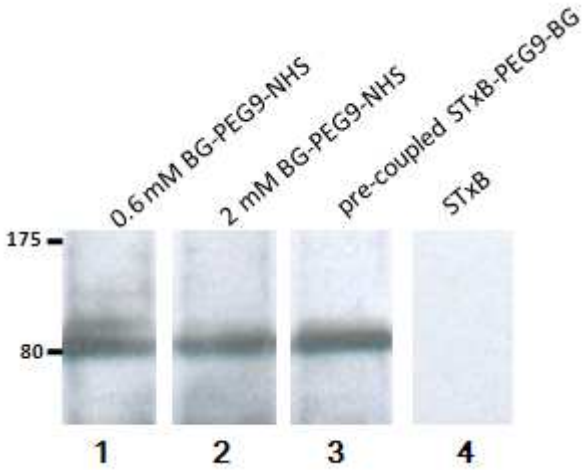


Figure 27. Cell surface modification. STxB was bound to pGalT-GFP-SNAP transfected cells at 4°C for 45min. Cells were then incubated on ice with different concentrations of BG-PEG$_9$-NHS(0.6 or 2mM), and then shifted to 37°C for 3 hours. After blocking and cell lysis, the fusion protein was immunoprecipitated with anti-GFP antibodies, and immunoprecipitated material was analyzed by Western blotting with antibodies against STxB. Pre-coupled STxB-PEG$_9$-BG was bound to cell surface and follow the same process in lane 3. Lane 4 represented the negative control by binding STxB to cell without BG-PEG$_9$-NHS surface labeling.

Identifier	Size (aa)	Mol. weight (Da)	Match group	Score	Matching peptides
					All
GalTGFP	366	39608	GalTGFP	584.69	11
SNAP	193	20239.3	SNAP	252.42	4
STxB	70	7793.8	STxB	46.55	1

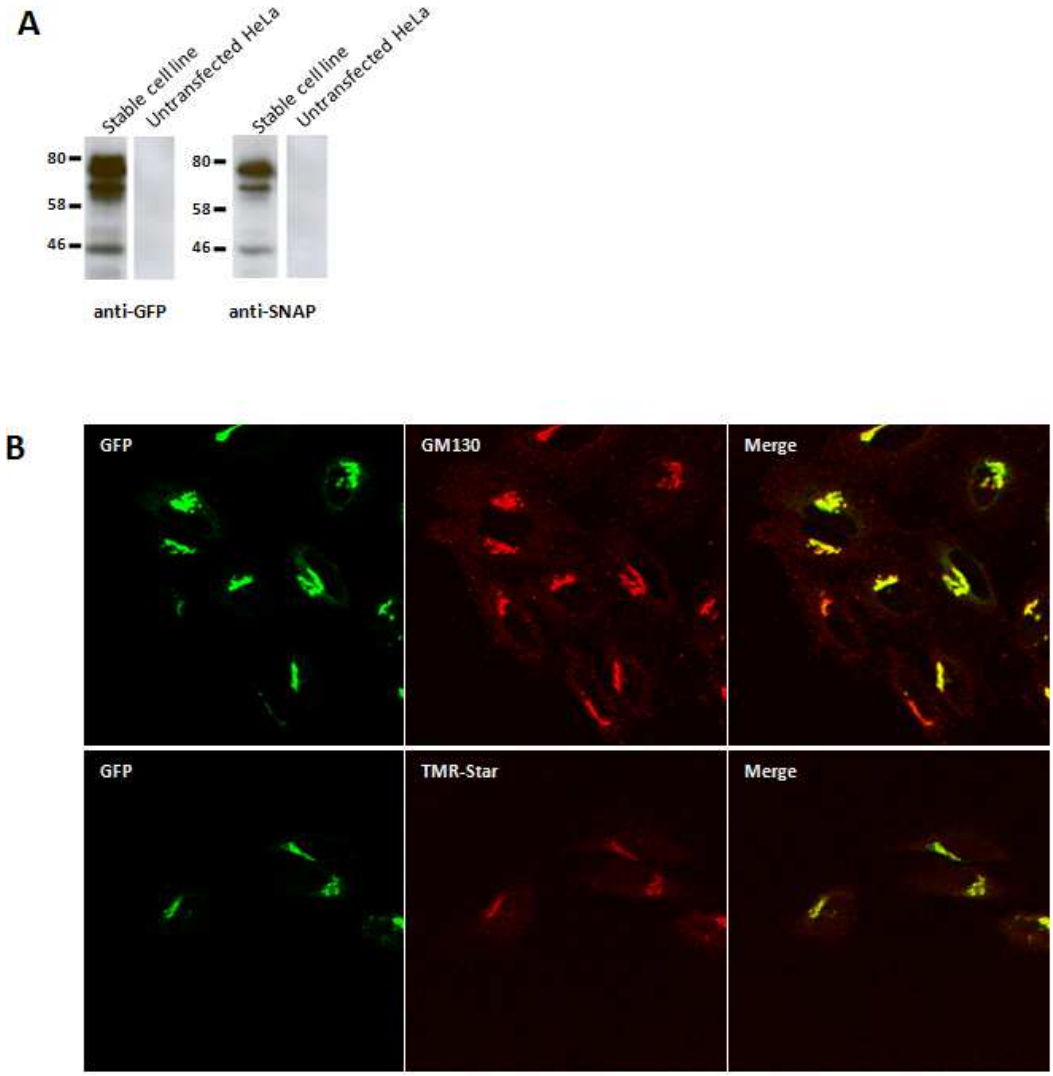
Table 8. LC-MS identification of STxB in cell surface modification assay. Peptides from STxB as well as from GalT-GFP-SNAP were identified.

1.3 Proteomics screening

In proteomics, large quantities of protein are needed for LC-MS. To fulfill this requirement, a stable cell line expressing GalT-GFP-SNAP was cloned. The expressing and localization of the chimeric protein were verified by Western blotting and IF analysis. SNAP and GFP colocalized, as expected (Figure 28A), GFP colocalized with GM130 (Figure 28B, upper panel), and TMR-Star displayed same pattern as GFP (Figure 28B, lower panel).

The use of antibodies for immunoprecipitation has disadvantages in proteomics.

Antibody fragments can cause signal saturation in LC-MS, thereby preventing the identification of unknown proteins. Traditionally, antibodies are immobilized onto solid surfaces to get around this problem. Here, we used a new immunoprecipitation method, termed GFP-Trap. GFP-Trap is a 13 kDa single domain antibody fragment that is covalently linked to agorose beads. This antibody fragment maintains high affinity for GFP, with a Kd of 0.59 nM. Due to the high density of antibody covering the bead surface, unspecific binding is strongly reduced. To perform proteomics, the stable GalT-GFP-SNAP expressing cell line was first incubated with BG-PEG9-NHS at 4°C for 1h. This cell surface modification was performed in culture dishes without detaching cells. The cells were then shifted to 37°C for 10 hours. We reasoned that such long incubation time would be favorable in the case of proteins that have a slow intracellular trafficking kinetics. As before, the blocking agent was used to quench remaining SNAP-tag activity, cells were lysed, lysates incubated with GFP-Trap, and immunoprecipitated proteins were run on a SDS-PAGE gel. Entire lanes were cut from the gel, proteins were cleaved with trypsin, and peptides released from the gel were identified by LC-MS. Here, the gel step was used not so much to separate proteins, but rather to make sure that they are optimally unfolded for trypsin cleavage.



Figuer 28. Verification of GalT-GFP-SNAP expression in stale cell line. (A)cell lysate from stable cell line was detected by Western blotting using anti-GFP and anti-SNAP antibody. (B)IF analysis showed colocalization of GFP with Golgi marker GM130(upper panel) and TMR-Star stained functional SNAP-tag which showed the same pattern as GFP(lower panel).

Contaminations cannot be avoided when using immunoprecipitation for protein purification, even with GFP-Trap. As mentioned in introduction, comparative proteomics is a powerful method to address this point. We therefore sought to develop a control condition: GalT-GFP-SNAP expressing cells that were treated exactly as described above, except for the fact that they were not incubated with BG-PEG$_9$-NHS. By comparing the proteins that are identified in both conditions, contamination can be eliminated to some extent. With this, 74 proteins were identified that were specific to the BG-PEG$_9$-NHS condition. Of these, 20 look like interesting candidates. They are listed in Table 9 and discussed in the discussion section. Several proteins were found multiple times in different experiments (3 times in total).

ID	Entrez Gene Name	Location	Type(s)	Peptide identified	Number of times found
21361181	ATPase, Na+/K+ transporting, alpha 1 polypeptide	Plasma Membrane	transporter	7	1
119574322	ATP-binding cassette, sub-family C (CFTR/MRP), member	Plasma Membrane	transporter	1	2
17978489	CD97 molecule	Plasma Membrane	G-protein coupled receptor	1	1
171543879	CD99 molecule	Plasma Membrane	other	2	3
5915662	integrin alpha 11 subunit precursor	Plasma Membrane	other	4	1
220141	integrin, alpha 3 (antigen CD49C, alpha 3 subunit of VLA-3 receptor)	Plasma Membrane	other	3	2
119617186	integrin, alpha 5 (fibronectin receptor, alpha polypeptide), isoform CRA_a	Plasma Membrane	transmembrane receptor	2	1
19743813	integrin, beta 1 (fibronectin receptor, beta polypeptide, antigen CD29 includes MDF2, MSK12)	Plasma Membrane	transmembrane receptor	5	1
46370049	major histocompatibility complex, class I, A	Plasma Membrane	transmembrane	3	1
110624774	mannose receptor, C type 2	Plasma Membrane	transmembrane	1	2
19387846	melanoma antigen family D, 2	Plasma Membrane	other	1	1
76879878	plexin B2	Plasma Membrane	other	5	1
33598950	podocalyxin-like	Plasma Membrane	kinase	3	2
15004317	solute carrier family 1 (neutral amino acid transporter), member 5	Plasma Membrane	transporter	1	2
4759112	solute carrier family 16, member 3 (monocarboxylic acid transporter 4)	Plasma Membrane	transporter	4	2
61744483	solute carrier family 3 (activators of dibasic and neutral amino acid transport), member 2	Plasma Membrane	transporter	6	2
16041779	solute carrier family 39 (zinc transporter), member 14	Plasma Membrane	transporter	1	2
24981008	solute carrier family 7 (cationic amino acid transporter, y+ system), member 5	Plasma Membrane	transporter	2	3
24475816	trans-2,3-enoyl-CoA reductase	Plasma Membrane	other	3	1
189458817	transferrin receptor (p90, CD71)	Plasma Membrane	transporter	11	1

Table 9. Protein candidates for retrograde transport from plasma membrane to Golgi apparatus.

2. Sulfation approach

The sulfation approach to discover retrograde transport cargo proteins is outlined in Figure 29A. Similar to SNAP-tag approach, a probe termed bSuPeR (biotinylated sulfation site peptide reagent) is used to tag plasma membrane proteins. bSuPeR bears three functional entities: a NHS group for covalent crosslinking to plasma membrane proteins, a sulfation-site peptide (EEPEYGE; substrate for sulfotransferase), and a biotin for protein purification. The bSuPeR tagged plasma membrane proteins are expected to traffic through different pathways. Only the ones that use the retrograde transport route are delivered to the TGN and are sulfated. If incubation is done in the presence of the sulfate isotope $^{35}SO_4^{2-}$, TGN-localized bSuPeR-tagged proteins will become radioactive upon modification by sulfotransferase. After cell lysis, bSuPeR-tagged proteins are purified with streptavidin. Purified membrane proteins are separated by SDS-PAGE followed by autoradiography to visualize ^{35}S tagged species that can then be identified by MS.

In this project, four different bSuPeRs variants were tested which were termed bSuPeR-3b, bSuPeR-carbamate, bSuPeR-ester, bSuPeR-R-ester (Figure 30).

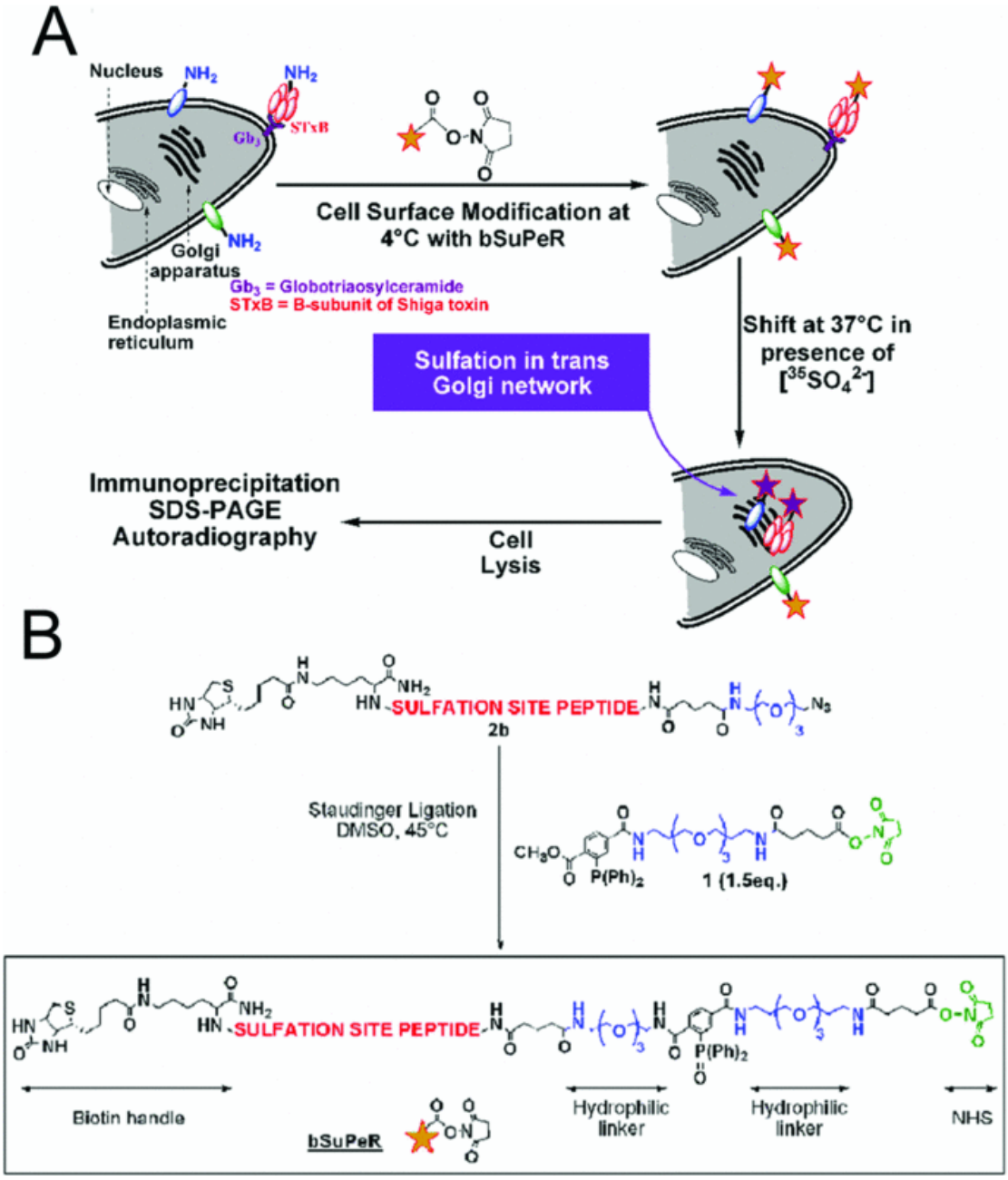


Figure 29. Scheme of sulfation approach(Annex1). (A) Strategy for sulfation approach: Plasma membrane proteins were modified at 4°C with bSuPeR. Cells were shifted to 37°C and bSuPeR-tagged proteins undergoing retrograde transport were modified by $^{35}SO_4^{2-}$ in TGN. Tagged proteins were immunoprecipitated by streptavidin. After SDS-PAGE, protein candidates were visualized by autoradiography through their radioactive sulfation-site peptide. LC-MS was used to identify these proteins. (B) Staudinger ligation ofr assembling phosphine compound with biotinylated azido compoud, giving bSuPeR-3b.

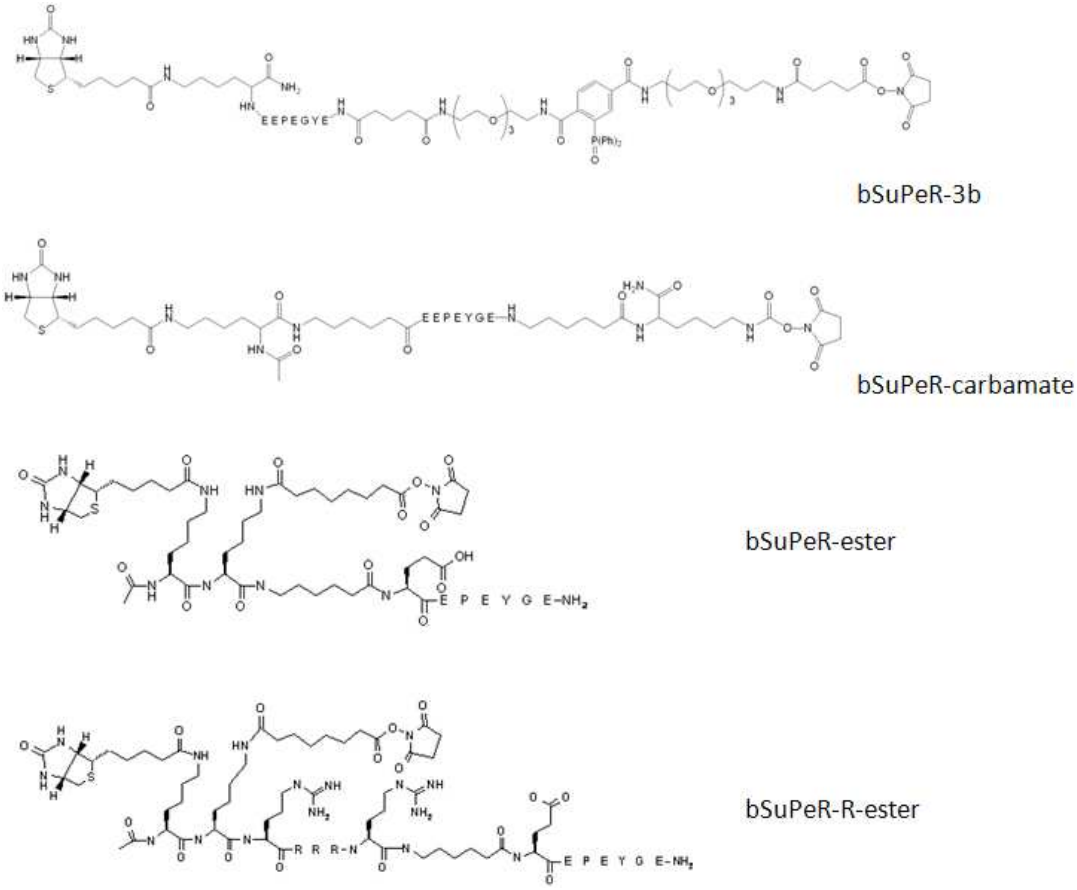


Figure 30. Four different bSuPeRs.

2.1 The synthesis of bSuPeRs

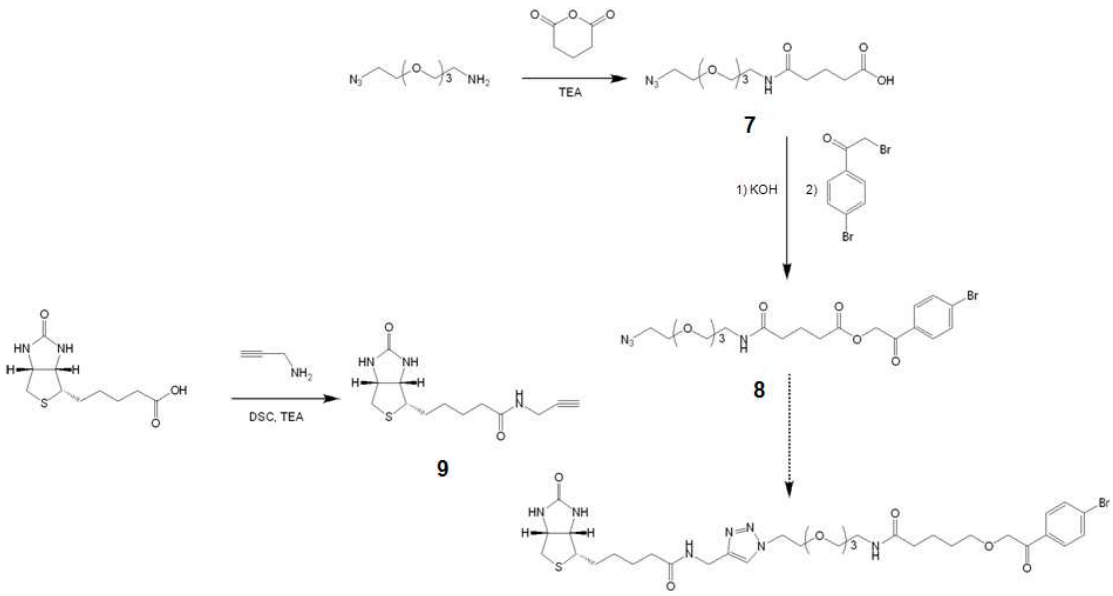
2.1.1 bSuPeR-3b synthesis

The synthesis of bSuPeR-3b has been described in our previous study [Annex 1] through Staudinger ligation (Figure 29B). Originally, bSuPeR-3b was designed to modify the cell surface in a two-step strategy. Firstly, a linker with NHS group is coupled to membrane proteins. Secondly a biotin-peptide group is coupled to the linker by Staudinger ligation (Saxon and Bertozzi 2000). An alternative strategy is to complete bSuPeR-3b synthesis by organic chemistry, and to modify cell surface in a one-step reaction.

2.1.2 Click reaction

To synthesize bSuPeR by Staudinger ligation is laborious. To simplify bSuPeR synthesis, the click reaction was introduced to link the NHS group to the peptide. One of the most popular click reactions is the azide alkyne Huisgen cycloaddition, named after Rolf Huisgen who first discovered it (Huisgen 1961). The reaction uses a Copper (Cu) catalyst and can proceed efficiently at room temperature. Two model compounds were synthesized to optimize reaction conditions following scheme 2. 2,4’-dibromoacetophenone was used to monitor the reaction by UV.

The two compounds were successfully obtained. In the mean time, bSuPeR-carbamate was also successfully synthesized using an even simpler strategy. We therefore discontinued the use of the click reaction to obtain bSuPeR.



Scheme 2. Verification of best condition for click reaction. Compound 8 and 9 were synthesized as model compound to verify the best condition for the azide alkyne Huisgen cycloaddition with a probe group 2,4’-dibromoacetophenone.

2.1.3 bSuPeR-carbamate synthesis

In collaboration with Dr. Melnyk, we could show that a peptide with a lysine residue can be converted into N-succinimidyl carbamate (NSC) derivatives by incubation with disuccinimidylcarbonate (DSC). For details on the synthesis of bSuPeR-carbamate, please refer to Annex 2.

2.2 Experimental setup

Our previous study has provided proof of concept that STxB can be cell surface modified by bSuPeR-3b and subsequently be tagged with radioactive sulfate by tyrosine sulfotransferase on its arrival in the TGN [Annex1]. This is a strong indication that the sulfation approach is feasible for proteomics of retrograde transport cargo proteins. However, during this study, we encountered the persistent problem that the sulfation signal for cell surface bSuPeR-3b modified STxB was remarkably lower than the one observed with *in vitro* bSuPeR-3b modified STxB. A series of experiments was carried out to analyze and to solve this problem.

2.2.1 Estimation of the efficiency for cell surface modification

The most direct and obvious explanation for the sulfation signal difference between *in vitro* and cell surface bSuPeR-3b modified STxB is that the cell surface modification is inefficient. We therefore designed a surface modification assay to characterize and quantitatively measure cell surface modification efficiency with bSuPeR-3b.

STxB was bound to HeLa cells at 4°C. The cells were then incubated at 4°C for 45min with bSuPeR-3b(4mM in PBS), lysed after washing, followed by streptavidin and/or antibody immunoprecipitation. Streptavidin immunoprecipitation captured all bSuPeR-3b tagged proteins including modified STxB via the biotin group of

bSuPeR-3b. The anti-STxB antibody immunoprecipitation pulled down the entire amount of cell associated STxB, modified or not by bSuPeR-3b. Tris-tricine gel electrophoresis and Western blotting using anti-STxB antibodies were performed (Figure 31). The upper bands (single star) represented bSuPeR-3b modified STxB and the lower bands (double stars) represent unmodified STxB. The comparison of upper and lower band intensities suggested that only a small proportion of STxB was modified by bSuPeR-3b. The upper band was not found in lane 2 since the antibody immunoprecipitation was performed after streptavidin had pulled down all the bSuPeR-3b tagged STxB, while lane 3 and 4 represented the two kinds of pulling down assays individually. The lower bands in streptavidin immunoprecipitation were due to the pentamer structure of STxB. During surface modification, not every monomer of STxB was modified. However, the unmodified monomers were also pulled down within the pentamer of STxB.

Here we have shown that cell surface modification by bSuPeR-3b is not complete. On the basis of the premise that the percentage of modified over non modified STxB is equal to the percentage of modified over non modified plasma membrane proteins, it become apparent that the reaction must be optimized.

2.2.2 bSuPeR-carbamate characterization

bSuPeR-carbamate molecules that were synthesized in collaboration with Dr. Melnyk were characterized by sulfation assay and IF analysis. bSuPeR-carbamate tagged STxB was properly sulfated and correctly located in Golgi apparatus after endocytosis (for detailed data, see Annex 2). As mentioned in Annex 2, bSuPeR-carbamate can easily form a cyclic peptide species by reaction between the phenol group of tyrosine and the NHS group, which should largely reduce cell surface labeling efficiency. With the help of Dr. Melnyk, a new NHS ester generation of bSuPeR was synthesized that do not form cyclic peptides. These bSuPeR-esters will be described in 2.2.4.

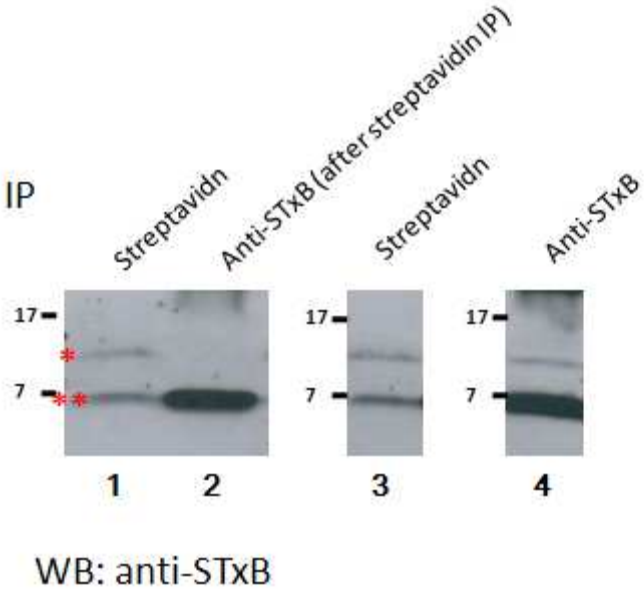


Figure 31. Estimation of the efficiency for cell surface modification by bSuPeR-3b. STxB(1μM) was bound to HeLa cells at 4°C. The cells were then incubated at 4°C for 45min with bSuPeR-3b(4mM in PBS). Cell lysate was immunoprecipitated by streptavidin or anti-STxB antibody(13C4) in lane 3 and 4 respectively. In lane 2, immnoprecipitation of anti-STxB antibody was performed using the supernatant of streptavidin IP in lane 1. Another anti-STxB antibody was used in Western blotting to visualize STxB. Upper band represented bSuPeR-3b modified STxB while lower band was for unmodified one.

2.2.3 An important element for efficient cell surface modification: reaction pH

The coupling reaction between the NHS group and amine group is highly dependent on pH. The reaction conditions are often buffered, and the pH range is controlled from 7 to 9. In bSuPeR-3b surface labeling, bSuPeR-3b solubilized in high concentration in DMSO is diluted to PBS buffer to a final concentration of 4 mM. Remarkably, we found that the addition of bSuPeR-3b changed the pH of PBS buffer from 7.4 to 6.3, due to the four glutamic acids in bSuPeR-3b. A surface modification assay was performed to verify the effect of pH in surface labeling. PBS with 4 mM final concentration of bSuPeR-3b in pH 7.4 or pH 6.3 was incubated with cells onto which STxB was bound. Following cell lysis, STxB was immunoprecipitated with anti-STxB antibody(13C4). After a tris-tricine gel electrophoresis, Western blotting using antibodies against STxB was performed. The upper band, representing bSuPeR-3b labeled STxB, was much weaker when modification was performed at pH 6.3, compared to the pH 7.4 condition (Figure 32). This indicated that a low pH was one reason for low efficiency of surface modification.

Even though we found that choosing the right pH would increase the cell surface labeling efficiency, the percentage of bSuPeR-3b labeled STxB was still too low, even at pH 7.4. This suggested that additional parameters needed to be taken into consideration for the optimization of surface modification efficiency with bSuPeR-3b.

2.2.4 A key element for efficient cell surface modification: the peptide charge

Previous studies have showed that charge plays an important role in membrane surface interactions (Brac 1984; Zolla, Lupidi et al. 1990) as well as protein-protein interactions (Graves 1988; Haugland and You 2008). We hypothesized that the reason for the low efficiency of cell surface modification with bSuPeR-3b is possibly linked to the highly negative charge of the sulfation-site peptide (EEPEYGE).

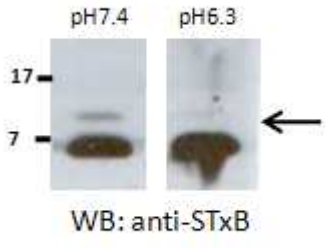


Figure 32. Cell surface modification in different pH. STxB was bound to HeLa cells at 4°C. The cells were then incubated at 4°C for 45min with bSuPeR-3b(4mM in PBS). In pH7.4 condition, the pH of bSuPeR-3b PBS solution was adjusted to 7.4 before incubation. Cell lysate was immunoprecipitated by anti-STxB antibody(13C4). Another anti-STxB antibody was used in Western blotting to visualize STxB.

The peptide contains 4 glutamic acids, and its pI is 3.58. The negatively charged bSuPeR-3b will possibly be repelled from the negativly charged cell membrane.

To study the effect of charge on cell surface modification, two different bSuPeR versions were synthesized by Dr. Melnyk's team, termed bSuPeR-ester and bSuPeR-R-ester (Figure 30). In bSuPeR-R 5 arginines were added to the peptide to neutralize the 4 glutamic acids. This changed the pI of the peptide from 3.58 to 8.74. During surface labeling in pH 7.4, bSuPeR-ester is expected to be negatively charged, and bSuPeR-R-ester positively.

bSuPeR-ester and bSuPeR-R-ester were coupled to STxB and characterized by IF. Anti-STxB or anti-biotin antibodies were used to detect total STxB or modified STxB, respectively. The colocalization of biotin and STxB with GM130 showed that the modification did not affect the transport characteristics of STxB (Figure 33).

To analyze the effect of charge on surface modification, the sulfation assay was performed using STxB as a model. STxB was bound at 4°C to the plasma membrane of HeLa cells, which were then incubated still at 4°C with bSuPeR-ester or bSuPeR-R-ester(4mM) in PBS at pH 7.4 for 1h. The genetic fusion protein STxB-Sulf2 was used as a positive control to provid the maximum signal. The cells were then shifted for 2 hours to 37°C in the presence of radioactive sulfate, lysed, immunoprecipitated by anti-STxB antibody, and sulfated STxB were revealed by Tris-Tricine gel electrophoresis and autoradiography. Very strikingly, bSuPeR-R-ester dramatically increased STxB sulfation compared with bSuPeR-ester (Figure 34A). The total sulfation signal for bSuPeR-R-ester modified STxB was 41% of that observed with STxB-Sulf2. Since the latter has two sulfation sites, and bSuPeR-R-ester only one, it can be assumed that most of the cell associated STxB had reacted with bSuPeR-R-ester during the cell surface modification reaction.

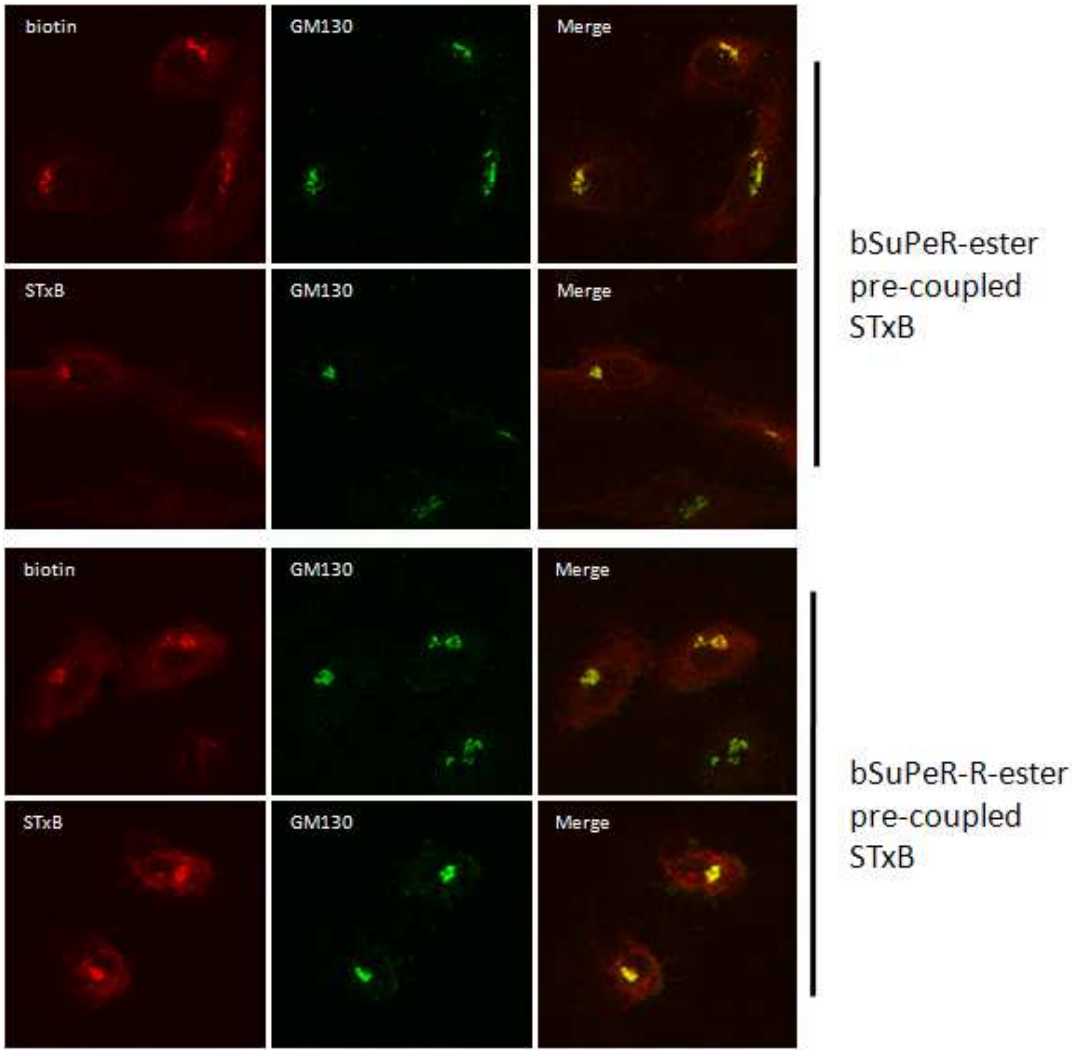


Figure 33. IF analysis for bSuPeR-ester and bSuPeR-R-ester pre-coupled STxB. Pre-coupled STxB(1μM) was incubated with HeLa cells at 4°C. After washing, cells were then shifted to 37°C for 1 h. Cells were fixed and permeabilized. Anti-biotin and anti-STxB antibody was used in IF analysis to detect the localization of coupled STxB and total STxB.

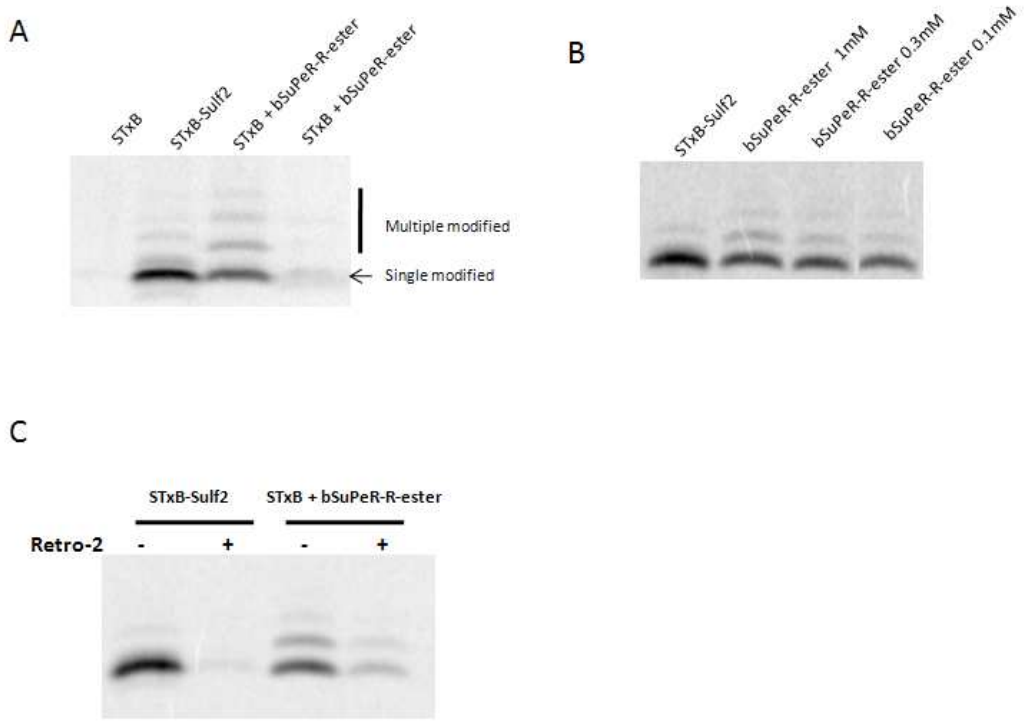


Figure 34. Sulfation assay of bSuPeR-ester and bSuPeR-R-ester cell surface modification. (A) After STxB or STxB-Sulf2 were bound to HeLa cells at 4°C, cells were incubated at 4°C for 45min with bSuPeR-ester or bSuPeR-R-ester(4mM in PBS) in pH7.4. The cells were then shifted for 2 hours to 37°C in the presence of radioactive sulfate. STxB were immunoprecipitated by anti-STxB antibody and visulized by autoradiography. (B) Different concentrations of bSuPeR-R-ester(1, 0.3 and 0.1 mM) were used to modified cell surface bound STxB, followed sulfation assay. (C) Retrograde transport inhibitor Retro-2(20 μM) was used to verify the transport specificity of the of bSuPeR-R-ester surface modified STxB.

Different concentrations of bSuPeR-R-ester (0.1, 0.3, or 1 mM final concentrations) were used to determine the most efficient conditions for cell surface modification. The signal started decreasing only in the 0.1 mM condition, which demonstrated that 0.3 mM of bSuPeR-R-ester was sufficient to have opitmal cell surface modification (Figure 34B). The small molecule retrograde transport inhibitor Retro-2 was used to verify the transport specificity of the of bSuPeR-R-ester modified STxB. Retro-2(20 μM) was added to bSuPeR-R-ester surface-modified cells to which STxB had been bound, or to cells with STxB-Sulf2. In both cases, a strong decrease of sulfation was observed (Figure 34C), demonstrating the functionality of STxB in both conditions.

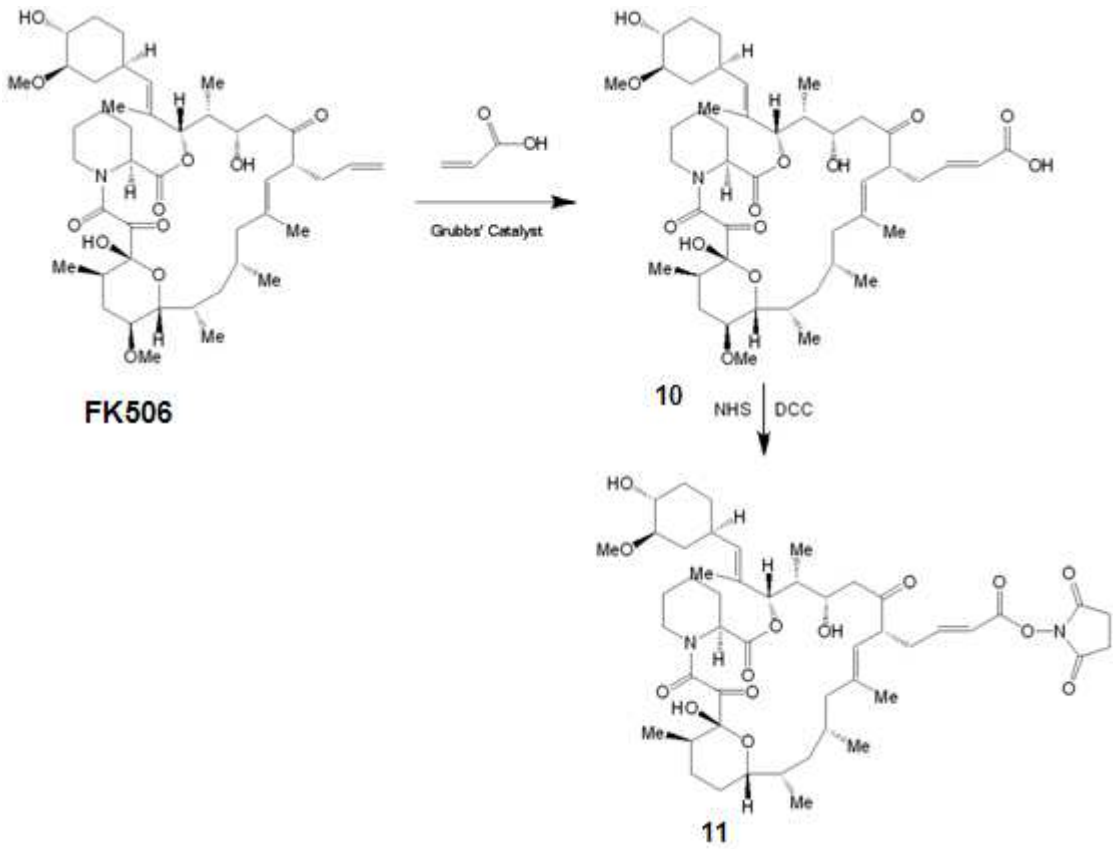
3. FKBP approach

FKBP12 has a strong affinity for FK506 with a Kd of 0.6 nM. This strong interaction could possibly be exploited for proteomics to screen retrograde transport cargo proteins. A similar bait and trap principle as the one of the SNAP-tag approach applies here. In this strategy, FK506 is chemically linked to the NHS group. A fusion protein consisting of FKBP12, GFP and a resident Golgi protein is expressed in cells to localize FKBP12 in the Golgi apparatus. Plasma membrane proteins are modified with FK506-NHS, and those that undergo retrograde transport will be put in interaction with the Golgi-localized FKBP12. After cell lysis, retrograde proteins can be co-immunoprecipitated with anti-GFP antibodies and identified by LC-MS.

3.1 FK506-NHS synthesis

Rapamycin and FK506 both bind FKBP12 with high affinity. To preserve the binding capacity for FKBP12, the NHS group must be introduced into the effector domain. FK506, but not Rapamycin carries a unique olefin moiety (C39-C40) in the effector domain, and we therefore choose FK506 to generate the NHS probe.

Previous studies on the synthesis of a FK506 dimer (Schreiber 1997) indicated a synthesis path for the generation of FK506 derivatives by olefin metathesis. Grubbs' ruthenium metathesis catalyst 1 was used to catalyze the metathesis between FK506 and acrylic acid (Scheme 3) to introduce a carboxyl group into FK506. The carboxyl of purified FK506-COOH was further activated by NHS to finally obtain FK506-NHS.



Scheme 3. Synthesis of FK506-NHS.

3.2 Verification strategy

3.2.1 In vitro binding

a). STxB coupling to FK506-NHS

STxB in PBS (150 µM) was incubated at RT for 1h with FK506-NHS in the molar ratios of 1:1, 1:3, or 1:27. The samples were analyzed by Maldi-TOF (Figure 35). Unmodified STxB was detected as a peak at 7792 Da. During the reaction, the NHS group is released, which leads to an 832 Da increase in size on STxB. Indeed, single modified STxB at 8624 Da was observed. The molar ratio 1:3 showed optimal modification of STxB.

b). Binding assay in solution

Purified FKBP12 protein is commercially available. A binding assay between STxB-FK506 and purified FKBP12 was performed in solution. STxB-FK506 and FKBP12 were mixed in PBS buffer in 1:1 molar ratio(both in 50µM final concentration). The solution was incubated for 2 hours at 37°C. Proteins were immunoprecipitated by anti-STxB antibodies. If FKBP12 bound to STxB-FK506, it should be co-immunoprecipitated. This was tested by Western blotting using anti-FKBP12 antibodies (Figure 36A). In the negative control (lane 2), unmodified STxB instead of STxB-FK506 was incubated with FKBP12. The weak band in lane 2 was due to the unspecific binding of FKBP during immunoprecipitation. These experiments showed that FKBP12 could bind to FK506-modified STxB and be co-immunoprecipitated.

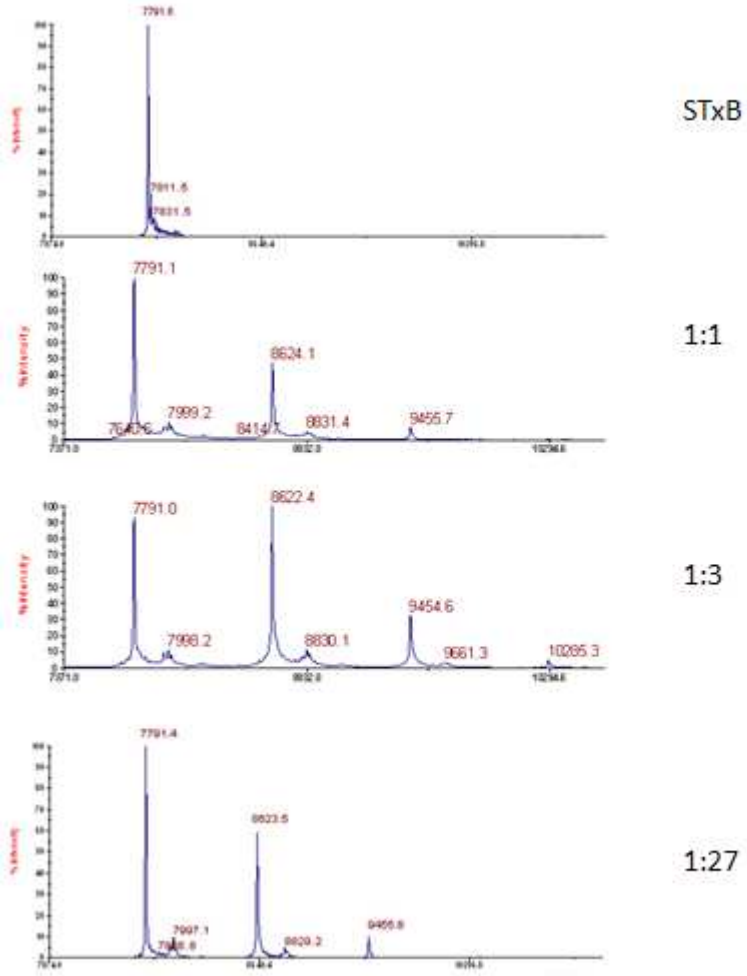


Figure 35. STxB modification by FK506-NHS in different concentrations.

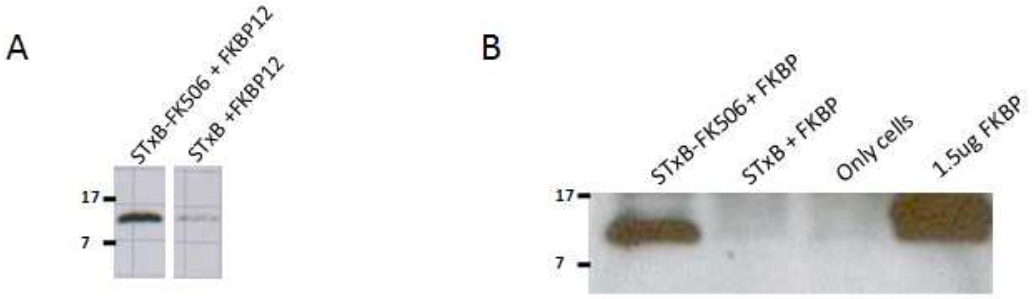


Figure 36. In vitro binding between FKBP12 and STxB-FK506. (A) STxB-FK506/STxB was mixed with FKBP12 in molar ratio 1:1 in 50ul final concentration. The solution was incubated for 2 hours at 37°C. Proteins were immunoprecipitated by anti-STxB antibodies. Western blotting was performed using anti-FKBP12 antibodies. (B) STxB-FK506 was bound at 4°C to HeLa cells. FKBP12 was then added at 4°C and incubated for 1h. Cells were lysed and STxB was immunoprecipitated. SDS-PAGE and Western blotting were performed using anti-FKBP12 antibodies to detect FKBP12.

c). Binding assay on cell surface

To further validate the binding between FKBP12 and FK506-labeled STxB in the cellular context, STxB-FK506 was bound at 4°C to HeLa cells. As mentioned before, STxB tightly bound to cells under these conditions without being taken up. Purified FKBP12 was then added at 4°C and incubated for 1h. FKBP12 does not have a natural receptor on HeLa cells, and binding could only occur via STxB-FK506. After washing away unbound FKBP12, cells were lysed and STxB was immunoprecipitated using appropriate antibodies. SDS-PAGE and Western blotting were performed using anti-FKBP12 antibodies to detect FKBP12 that would have been pulled down with STxB-FK506 (Figure 36B). FKBP12 immunoreactivity could clearly be detected when STxB-FK506 and FKBP12 had been present (lane 1). Purified FKBP12 was loaded as a reference (lane 4). Two control conditions were used: non modified STxB instead of STxB-FK506 (lane 2); absence of FKBP12 (lane 3).

3.2.2 Localization of FKBP12 to the Golgi apparatus

To localize FKBP12 to the Golgi apparatus, the protein was fused to a 17-residue transmembrane domain sequence of β-Galactoside α2,6-sialyltransferase (ST). ST is a type II integral membrane protein of the Golgi apparatus involved in the sialylation of N-linked glycans. Previous studies have established that the 17-residue transmembrane domain of ST is sufficient for Golgi retention (Wong, Low et al. 1992). The pCMV-ST-FKBP12-GFP plasmid was kindly provided by Perez F. The expression and localization of the fusion protein ST-FKBP12-GFP (44.39kD) were verified by Western blotting using anti-GFP antibodies (Figure 37A), and IF analysis of transiently transfected cells (Figure 37B). Our anti-FKBP12 antibody did not function, neither in Western blotting nor by IF. Of note, the transfection efficiency was not very high (Figure 37B).

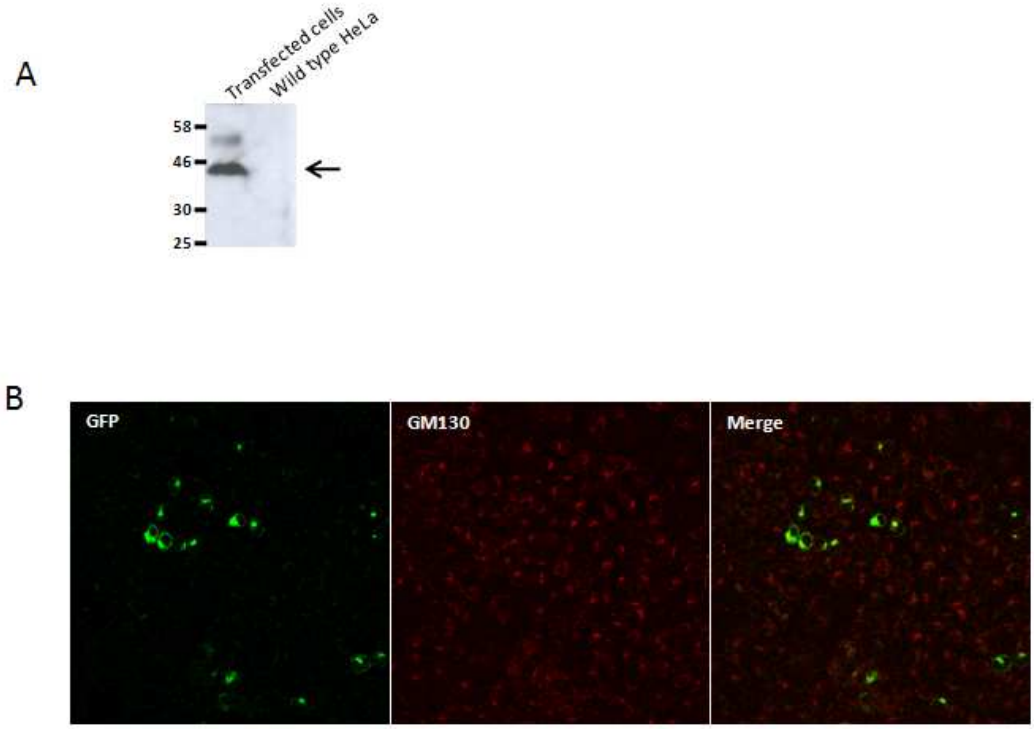


Figure 37. Verification of ST-FKBP12-GFP expression. (A)Western blotting using anti-GFP antibodies to detect ST-FKBP12-GFP (44.39kD) in transfected cell lysate. (B) IF analysis showed overal expression level of ST-FKBP12-GFP.

3.2.3 Functional validation

STxB-FK506 or STxB were incubated for 3 hours at 37°C with HeLa cells that had been transiently transfected with pCMV-ST-FKBP12-GFP. To prevent FKBP12 binding to FK506 after cell lysis, additional FK506 (200 μM) was added to the lysis buffer as a competitive inhibitor. Proteins in the cell lysates were immunoprecipitated with anti-STxB antibodies, and Western blotting with anti-GFP antibodies was used to search for co-immunoprecipitated ST-FKBP12-GFP. As shown in Figure 38, GFP immunoreactivity could be detected in conditions in which ST-FKBP12-GFP and STxB-FK506 had been present both (lane 1), but not when unmodified STxB had been used (lane 2).

These experiments demonstrated that STxB-FK506 was captured in Golgi apparatus by ST-FKBP12-GFP, and that the non-covalent bond formed between FKBP12 and FK506 was stable during immunoprecipitation.

4. Streptavidin approach

Shiga toxin travels the retrograde transport route all the way from the cell surface to the ER from where the catalytic A subunit is translocated to the cytosol. The proteomics strategies that were described above can screen proteins taking first part of this pathway: retrograde trafficking to the Golgi. The streptavidin approach proposes a potential proteomics format for endogenous proteins that reach the ER, as Shiga toxin.

The bait and trap model is again applied here. Streptavidin is fused to a resident ER protein to function as a trap, and biotin is used as a bait to label plasma membrane proteins. Streptavidin will capture biotin-tagged plasma membrane proteins in the ER once they arrive by retrograde transport.

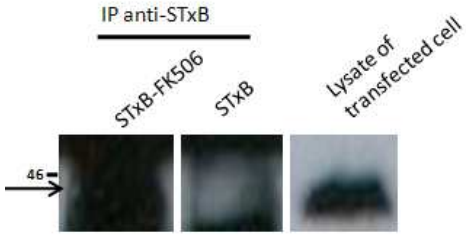


Figure 38. Capturing assay in Golgi apparatus. STxB-FK506 or STxB were incubated for 3 hours at 37°C with HeLa cells that had been transiently transfected with pCMV-ST-FKBP12-GFP. The cell lysates were immunoprecipitated with anti-STxB antibodies, and Western blotting with anti-GFP antibodies was used to search for co-immunoprecipitated ST-FKBP12-GFP. The third lane was loaded by transfected cell lysate without any treatment as the reference for ST-FKBP12-GFP.

To localize streptavidin in the ER, a short peptide termed Iip33 was fused to the N terminus of streptavidin. Iip33 is an ER resident polypeptide generated from human invariant chain (Ii) (Schutze, Peterson et al. 1994). The plasmid pCMV-Iip33-Strep-HA was kindly provided by Perez F.

The plasmid was transfected into HeLa cells and the expression of Iip33-strep-HA was tested by Western blotting and in immunofluorescence experiments. Our available antibodies did not allow us to obtain satisfactory results. The antibodies against Iip33 and streptavidin showed no bands on Western blots, and the anti-HA antibody gave strong non-specific labeling. Several commercially available antibodies were tested, without success at this stage. The immunofluorescence results were equally disappointing. Since other approaches gave more encouraging results, this strategy was temporarily put on hold. However, with proper antibodies, the streptavidin approach should also be a powerful way to study the retrograde route in all the way to the ER.

MATERIALS AND METHODS

1. Chemistry part

Solvents are purified as below:

- Dimethylformamide, acetone and methanol are dried by molecular sieve
- Dichloromethane is distilled by P_2O_5

The other solvents (ethyl acetate, cyclohexane…) with a "pure for synthesis" are utilized without additional purification.

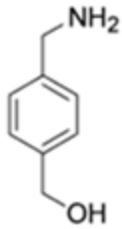
Thin layer chromatography (TLC) is performed on plastic plate supported Silica Gel 60 F254 (E.Merck, layer thickness 0.2mm). They are visualized under a UV lamp or by dipping in a 10% ethanol solution of phosphomolybdic acid and/or a 5% ethanol solution of concentrated sulfuric acid and/or in a solution of ninhydrin followed by ignition (when amines are present in the molecules). TLCs are used to track the progress of a reaction or to control the purity of a compound.

Flash column chromatography is performed on silica column (MERCK 230-600 mesh). The weight of silica is about 30 to 50 times of the weight of product to purify. The composition and proportions of the elution are the same as the product is separated in TCL with an Rf of 0.2-0.3.

Elctro spray ionization mass spectra (ESI-MS) are acquired with a quadruple instrument with WATERS ZQ2000.

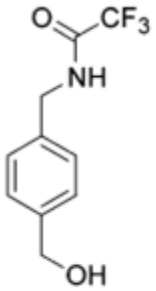
^{1}H NMR spectra is recorded on a Bruker AC 300 MHz spectrometer. Chemical shifts are reported in δ values relative to tetramethylsilane or DMSO-d6.

4-(aminomethyl)-benzyl alcohol [2]



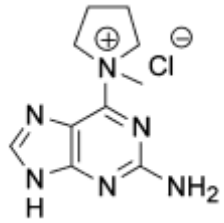
Commercially available 4-(cyanomethyl)benzoic acid (16.1 g, 0.1 mol) was dissolved in anhydrous THF (150 mL), and $LiAlH_4$ (15.2 g, 0.4 mol) was added in several portions at 0°C under N_2. After addition, the temperature was allowed to raise to room temperature, the reaction was stirred for 5 min, and then heated to reflux and stirred for another 5 h. LCMS and TLC using dichloromethane/methanol (10:1) showed the raw material was consumed. The reaction was cooled down, and quenched by the slow addition of ice water at 0-5°C. The mixture was treated with aqueous NaOH (10%, 50 mL), and then extracted with ethyl acetate (5 x 40 mL). The combined organic layers were washed twice with brine, dried over $MgSO_4$ and evaporated to dryness in vacuo. The crude compound **2** (10.3 g, 75%) was used without purification. Rf:0. 27; ^{1}H NMR (DMSO-d6) δ ppm 3.55 (s, 2H), 4.32 (s, 2H), 4.98 (bs, 1H), 7.12 (m, 4H); MS (ES^+) for $C_8H_{11}NO$: m/z 138 as $[M+H]^+$ peak product.

trifluoro-N-(4-(hydroxymethyl)benzyl)acetamide [3]



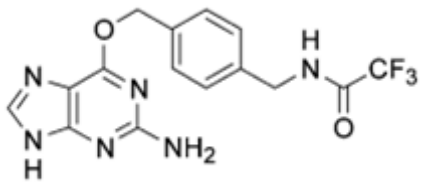
4.17 g (29.39 mmol) trifluoroacetic acid ethyl ester was added dropwise to a solution of 3.36 g (24.49 mmol) of compound 2 and 3.0 mL triethylamine in 30 mL dry methanol. The reaction mixture was stirred for 1 h and diluted with 50 mL ethyl acetate and 50 mL water. The aqueous layer was extracted with ethyl acetate and the combined organic layers were washed with saturated NaCl and dried over $MgSO_4$. After removal of the solvents in vacuo, the crude product was purified by flash chromatography (ethyl acetate/petrol-ether, 1/2) to yield compound **3** (4.37 g, 78%). Rf=0.28; ^{1}H NMR ($CDCl_3$) δ ppm 1:74 (br s, 1H), 4.53 (d, J = 5.8 Hz, 2H), 4.71 (s, 2H,), 6.60 (br s, 1H), 7.30 (m, 4H); MS (ES^+) for $C_{10}H_{10}F_3NO_2$: m/z 234 as $[M+H]^+$ peak product.

6-(pyrrolidinyl)-9H-purin-2-amine [4]



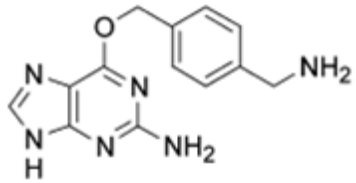
Commercially available 6-chloro-guanine (10.0 g, 66.1 mmol) was suspended in 350 mL DMF. The mixture was stirred for 30 minutes; then 1-methyl-pyrrolidin (22 mL, 200 mmol, 3 eq.) was added. After two days, 100 mL acetone was added to precipitate the product as a pale yellow solid. The solution was decanted and filtered. The solid was washed with acetone and dried in vacuo to obtain compound **4** (12.4 g, 79 %). Rf (CH_2Cl_2/MeOH, 1/1) = 0.03; ^{1}H NMR (DMSO-d6) δ ppm 13.38 (s, 1H), 10.84 (s, 1H), 8.35 (s, 1H), 7.11 (s, 2H), 4.60 (m, 2H), 3.96 (m, 2H), 3.65 (s, 3H), 2.24 (m, 2H), 2.05 (m, 2H); MS (ES^+) for $C_{10}H_{15}ClN_6$: m/z 220 as $[M+H]^+$ peak product.

N-(4-((2-amino-9H-purin-6-yloxy)methyl)benzyl)-2,2,2-trifluoroacetamide [5]



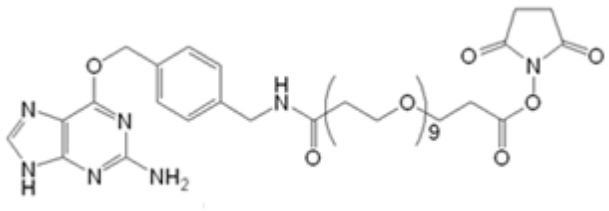
3.6 g (15.44 mmol) of compound **3** was dissolved in 15 mL dry DMF under argon atmosphere and 555 mg (23.15 mmol) NaH was added in three portions over 5 min. As much as 3.94 g (15.44 mmol) of compound 4 and 0.2 g (1.64 mmol) 4-(dimethylamino) pyridine(DMAP) were then added and the solution was stirred for 2 h at RT. One milliliter of water was added carefully to quench all excess NaH and the solvent was removed in vacuo. The crude product was dissolved in methanol, adsorbed on SiO_2 (10 g), and purified by flash chromatography (dichloromethane/methanol, 95/5) to obtain the compound **5** (4.9 g, 87%). Rf = 0.27; ^{1}H NMR (DMSO-d6) δ ppm 4:40 (d, J = 4.9 Hz, 2H), 5.46 (s, 2H), 6.29 (bs, 2H), 7.4 (m, 4H,), 7.81 (s, 1H), 10.01 (s, 1H), and 12.42 (s, 1H); MS (ES^+) for $C_{15}H_{13}F_3N_6O_2$: m/z 367 as$[M+H]^+$ peak product.

6-(aminomethylbenzyloxy)-9H-purin-2-amine [6]



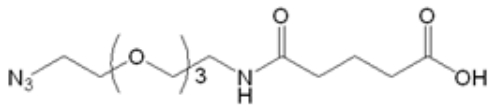
2.0 g (5.46 mmol) of compound 5 was dissolved under argon atmosphere in 10 ml dry methanol and treated with methylamine (33% in ethanol, 30 ml) at room temperature for 24 h. The solvents and excess methylamine were removed in vacuo. The residue was redissolved in methanol, adsorbed on SiO_2 (3.0 g), and purified by flash chromatography (dichloromethane/methanol, 5/1). Compound **6** (1.35 g, 92%) was obtained as a pale yellow solid. Rf = 0.03; ^{1}H NMR (DMSO-d6) δ ppm 7:82 (s, 1H), 7.44 (dd, 4H, J = 36.5 Hz), 6.27 (s, 2H), 5.45 (s, 2H), and 3.73 (s, 2H); MS (ES^+) for $C_{13}H_{14}N_6O$: m/z 271 as $[M+H]^+$ peak product.

2,5-dioxopyrrolidin-1-yl 1-(4-((2-amino-9H-purin-6-yloxy)methyl)phenyl)-3-oxo-6,9,12,15,18,21,24,27,30-nonaoxa-2-azatritriacontan-33-oate [1]



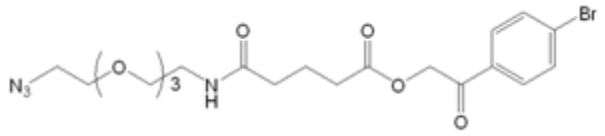
Compound 6 (78.2 mg, 0.29 mmol) was dissolved under argon in anhydrous DMF (10 mL), and was successively added to commercially available BS(PEG)9 (266.9 mg, 0.38 mmol), EDCI (78.3 mg, 0.41 mmol) and HOBT (55.4 mg, 0.41 mmol) solution at -5°C. The solution was stirred at room temperature for 6 h, TLC (MeOH/DCM, 1/2) and LC-MS showed the starting material was consumed. The solvent was removed under reduced pressure and the product was purified by prep.HPLC. The expected compound **1** was obtained (98 mg, 45%). Rf = 0.12; MS (ES^+) for $C_{39}H_{57}N_7O_{15}$: m/z 865 as $[M+H]^+$ peak product.

N-(11-azido-3,6,9-trioxatridecanyl)-glutarylamide [7]



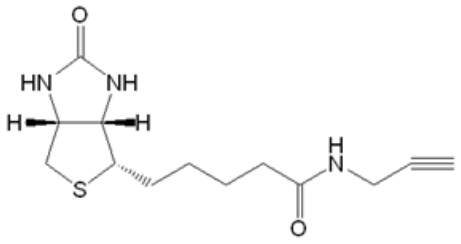
To a stirred solution of commercially available 11-azido-3,6,9-trioxatridecan-1-amine (1g, 4.6mmol) in dry THF (100mL) at room temperature TEA (715µL, 5mmol) and glutaric anhydride (575mg, 5mmol) were added under anhydrous conditions. The reaction mixture was stirred for 2 days and followed by TCL (C. hexane/Acetone, 5/5). The solvent was evaporated *in vacuo* and the crude product was purified by flash chromatography (C. hexane/Acetone/Acetic acid, 6/4/0,1) to give compound **7** (1.2g, 81%). Rf=0.26; ^{1}H NMR (300MHz, $CDCl_3$) δ ppm 1.94(m, 2H), 2.32(t, 2H, J=7.1), 2.41(t, 2H, J=6.8), 3.42(m, 5H), 3.55(t, 2H, J=4.4), 3.64(m, 10H), 5.29(s, 1H), 7.26(s, 2H); MS (ES^+) for $C_{13}H_{24}N_4O_6$: m/z 355 as $[M+Na]^+$ peak product.

2-(4-bromophenyl)-2-oxoethyl-1-azido-13-oxo-3,6,9-trioxa-12-azaheptadecanoate [8]



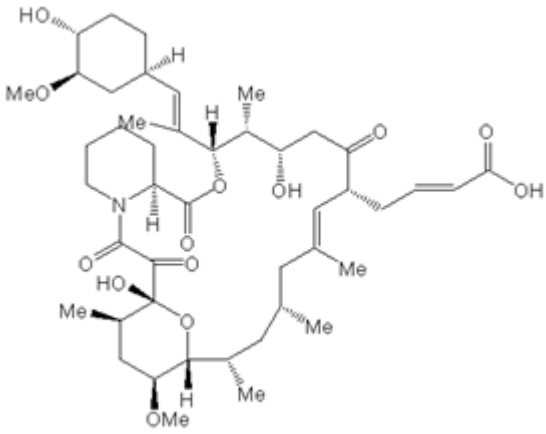
Compound 7 (220mg, 0.7mmol) was dissolved in methanol (10mL) with one drop of phenolphthalein (20mg/mL in methanol). A solution of KOH (100mg/mL) was added slowly until the pH 8.0. The solvent was removed and the crude product was dissolved in dry acetonitrile (10mL). Then, 2,4'-dibromoacetophenone (240mg, 0.85mmol) and dicyclohexyl-18-crown-6 (11mg, 0.04mmol) were added under anhydrous conditions. The reaction mixture was heated to 80°C for 15min until TLC analysis (C.hexane/Acetone, 5/5) showed that the reactant was consumed. Solvent was removed, and the crude product was purified on silica gel to provide compound **8** (260mg, 70%). Rf=0.34; ^{1}H NMR (300MHz, $CDCl_3$) δ ppm 2.16 (m, 3H), 2.34(m, 2H), 2.55(m, 2H), 3.37(m, 2H), 3.55(m, 2H), 3.64(m, 12H), 6.55(bs, 1H), 7.64(d, 2H, J=8.7Hz), 7.77(d, 2H, J=8.7Hz); MS (ES^+) for $C_{21}H_{29}BrN_4O_7$: m/z 557 as $[M+Na]^+$ peak product.

N-(propynyl)-biotinamide [9]



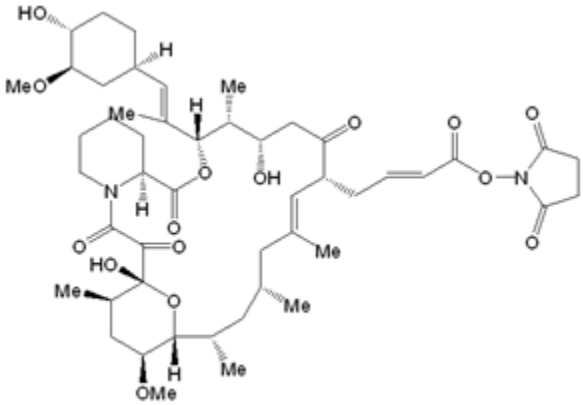
150mg (0.615mmol) of biotin was dissolved in 10mL of dry DMF under argon atmosphere. 118mg (0.46mmol) DSC and 260ul (1.84mmol) TEA were added. After 5h stirring at RT, 55μL(0.8mmol) propargylamine was added and the mixture was stirred at RT overnight. After the solvent was removed, the crude product was flash chromatographed on silica gel (MeOH/DCM, 1/9) to elute compound **9** (150mg, 87%). Rf=0.25; ^{1}H NMR (300MHz, $CDCl_3$) δ ppm 1.43(m, 2H), 1.63(m, 4H), 2.21(t, 2H, J=7.3Hz), 2.57(t, 1H, J=2.6Hz), 2.68(dd, 1H, J=5.0Hz), 2.90(m, 1H), 3.19(m, 1H), 3.93(d, 2H, J=2.70Hz)), 4.30(m, 1H), 4.48(m, 1H); MS (ES^+) for $C_{13}H_{19}N_3O_2S$: m/z 304 as $[M+Na]^+$ peak product.

FK506-COOH [10]



250mg (0.3mmol) of commercially available FK506 was dissolved in 100mL of dry DCM under argon atmosphere. 432μL (6mmol) acrylic acid was added. 20mg (0.024mmol) Grubbs' ruthenium metathesis catalyst 1 was added as catalyst for this Olefin metathesis. The reaction was followed by TLC and after 2days stirring at RT, the reactant FK506 was almost completely consumed. After removal of solvent, FK506-COOH was purified with a yield of 220.5mg(67%) on silica gel with an eluant mixture of MeOH/DCM(1:19); Rf=0.16; MS (ES^+) for $C_{44}H_{69}NO_{11}$: m/z 871 as $[M+Na]^+$ peak product.

FK506-NHS [11]



To a solution of compound 10 (20mg, 0.024mmol) in 3mL of dry DCM, 3.2mg (0.028mmol) NHS and 28μL (0.028mmol) DCC were added in anhydrous conditions. The reaction mixture was stirred at RT for 24 hours. The reaction was followed by TLC (MeOH/DCM:1/19). The solvent was removed and the crude product was purified by chromatography on silica gel (MeOH/DCM:1:19). 10mg(45%) of compound 11 was obtained; Rf=0.23; MS (ES^+) for $C_{45}H_{69}NO_{13}$: m/z 968 as$[M+Na]^+$ peak product.

2. Biology part.

Cell culture

HeLa cells were maintained in complete Dulbecco's modified Eagle's medium(DMEM Invtrogen) supplemented with 10% fetal calf serum, 5 mM L-glutamine, 100 unites/mL of penicillin and 100µg/mL of streptomycin in 5%CO_2 at 37°C in incubator(Water-jacketed Incubator, Forma Scientific)

Sulfation assay

Sulfation assay was performed as described (Mallard, Antony et al. 1998) using STxB, different peptide tagged STxB and STxB-$Sulf_2$ as a positive control (Johannes, Tenza et al. 1997). The proteins were incubated with 300,000 HeLa cells at 4°C in sulfate free Dulbecco's Modified Eagle Medium (DMEM, Invitrogen) during 45min. Cells were washed with sulfate free DMEM. After incubating with 480 µCi/mL radioactive sulfate ($^{35}SO_4^{2-}$, Perkin Elmer) in sulfate free DMEM at 37°C for 2h, cells were lysed in 900µL of RIPA buffer (150mM NaCl, 1% NP-40, 0.5% deoxycholic acid, 0.1% SDS, 50mM Tris pH 8.0) and 9 µL of protease-inhibitor cocktail (1mg/mL aprotinin, 1 mg/mL leupeptin, 1mg/mL antipain, 1mg/mL pepstatin, 1 M benzamidine, 40 mg/mL phenylmethanesulphonyl fluoride in DMSO). The cell lysate was precipitated with protein G-sepharose (GE Healthcare) with anti-STxB antibody 13C4 murine anti-STxB antibody (hybridoma from ATCC (Smith, Carvalho et al. 2006)). Precipitated proteins were analyzed by tris-tricine gel electrophoresis. Sulfated proteins were detected by autoradiography (STORM860, Molecular Dynamics).

Immunofluorescence analysis

200,000 HeLa cells were placed on coverslip in 24-well test plate(TPP) and if necessary incubated with 2 µM STxB or tagged STxB for 45 min at 4°C in PBS++ (PBS containing 0.9 mM CaCl2, 0.52 mM MgCl2 and 0.16 mM $MgSO_4$) and then shifted to 37°C for 1h (Water-jacketed Incubator, Forma Scientific). Cells were fixed with paraformaldehyde (300 µL of a 4% w/v solution in PBS), washed with PBS++

and then permeabilized with saponin solution (300µL, 0.2% w/vol saponin, 2% w/v BSA in PBS buffer). Each coverslip is incubated with 30ul saponin containing first antibody, washed with saponin solution, and then incubated with 30 µL saponin containing second antibody. Coverslips were mounted using 10% aqueous polyvinyl alcohol (Mowiol 4-88 from Fluka), 25% glycerol and 59% Tris-HCl 0.2M pH 8.5). The cells were analyzed by confocal microscopy (Leica Microsystems, Mannheim, Germany).

Procedure for antibody uptake assay

The antibodies(Roche) was prepared in 50 mM sodium phosphate/1 mM EDTA buffer, pH 7.5, at a concentration of 5mg/mL (31µM). Uncoupled reagent was removed by dialysis against 50 mM sodium phosphate/1 mM EDTA buffer, pH 7.5, with 10 kDa MWCO dialysis cassettes(Pierce). 400,000 MPR-GFP cells or HeLa cells were transiently transfected by pGalT-GFP-SNAP in 6-well tissue culture dish. 15 µg/ml anti-GFP/anti-GFP-BG in full medium was incubated with cells at 37° C for 4h. Cells were washed 3 times by full medium to remove external antibodies, and were blocked by SNAP-tag blocking reagent(1:200, NEB) in full medium at 37° C for 30min. Following 3 times full medium wash, 300µL lysis buffer(0.1%NP40 in TNE buffer(50mM Tris, pH7.5, 140mM NaCl, 5mM EDTA)) as well as protease-inhibitor cocktail was incubated with cells at 4°C for 30min. After centrifugation at 16100 × g for 10min, the supernatant was added to 50µL of Protein G Sepharose beads (Roche) and incubated with end-over-end rotation for overnight at 4°C. Precipitated proteins were analyzed by SDS-PAGE.

TMR-Star staining

TMR-Star(NEB) was dissolved in DMSO in concentration of 0.6mM. For staining SNAP-tag in cells, labeling solution was prepared by diluting TMR-Star(1:200) in cell culture medium. For 6-well plate, each well required 1mL labeling solution. After replacing the medium on the cells expressing SNAP-tag fusion protein with the

labeling medium, the cells were incubated at 37°C for 30min. Following 3 times washing by tissue culture medium, cells were incubated in fresh medium for another 30 min. The staining could be detected by IF analysis or directly visualized by fluorescence scanner on SDS gel.

Procedure for cell permeability test

400,000 HeLa cells, transiently transfected by pGalT-GFP-SNAP for 48h, were incubated at 4°C for 1h with 4mM BG-NHS (NEB), BG-PEG_3 (NEB), BG-PEG_9-NHS, or 10μM SNAP-Cell Block(NEB) in PBS++. The external compounds were then washed away, and cells were shifted to 37°C for 15 min. TMR-Star, a red fluorescent substrate of SNAP-tag, at 37°C for 30 min. TMR-Star staining was performed and cells were then lysed by lysis buffer(0.1%NP40 in TNE buffer(50mM Tris, pH7.5, 140mM NaCl, 5mM EDTA)) as well as protease-inhibitor cocktail at 4° C for 30min. Cell lysates were transferred to 1.5mL tube and after centrifugation at 16100 × g for 10min, the supernatant was subjected to SDS-PAGE. The red fluorescence of TMR-Star was directly visualized using fluorescence scanner (Typhoon 9400, GE).

Tris-tricine gel electrophoresis

5 mL gel contains 1.7ml 30% acrylamide,1.9% bis-acrylamide; 1.7ml 3M Tris-HCl with 0.3%SDS pH 8.45; 1.1 mL water; 0.7 mL glycerol; 50 μL 10% ammonium persulfate; 5 μL N,N,N',N'-tetramethylethylenediamine

SDS gel electrophoresis

7 mL gel contains 2.5mL 30%Acrylamide;1.9mL 1.5M Tris-HCl(pH 8.8); 20%SDS 75μL; 3mL water; 37.5μL 10% ammonium persulfate; 12.5μL N,N,N',N'-tetramethylethylenediamine.

Transient transfection

3µL of FuGENE (ROCHE) reagent was mixed with 97µL of DMEM at RT for 5min. 1µg of plasmid was added to mixture and incubate in RT for 30min. Transfection mixture were then added to culture medium of 300,000 cells prepared in 6-well culture dish. The transfection was finished in 6 h in 37°C. For better protein expression, cells were maintained in the incubation for 48h at 37°C.

Stable cell line establishment

pGalT-GFP-SNAP was transiently transfected to 3,000,000 HeLa cells by FuGENE for 48 h. Cells were split to three 9cm Petri dishes with dilution of 1/4, 1/10 and 1/100. 0.5mg/mL Geneticin(GIBCO) in full medium was incubated with cells to have the selection. Culture medium with 0.5mg/mL Geneticin was changed every 3 days for 2 weeks until cell colonies were visible. The colonies were picked by autoclaved small piece of Whatman paper that absorbed trypsin solution. Each piece of Whatman paper was incubated in full medium with 0.5mg/mL Geneticin in 24-well culture dish. The expression of protein was verified further by Western blotting and IF analysis.

DNA amplification by PCR

For a 50 µL reaction, 10 µL 5X reaction buffer, 28µL water, 1µL 10mM dNTPs, 5µL each 10 µM primer, 0.5 µL phusion enzyme(Finnzymes) and 1µL of template DNA were mixed on ice. The PCR was carried out by Mastercycler gradient(Eppendorf). The cycling followed denaturing 10 sec at 98°C, annealing 30sec at Tm+3°C and extending at 72°C for 2min. After 30 cycles, mixture was kept in 4°C.

DNA digestion and ligation

30µL digestion mixture consisted of 1µg of plasmid or DNA fragment, 3µL of desired

10x buffer(NEB), 0.3µL of BSA(NEB) if necessary, 1µL of restriction enzyme, and 24µL of water. The mixture was incubated in 37°C for overnight. For PmlI with blunt cutting, alkaline phosphatase, Calf Intestinal (CIP)(NEB) was added to digestion mixture and incubated at RT for 1h. DNA was purified using QIAquick PCR Purification Kit(QIAGEN). For 20µL DNA ligation reaction, 2µL 10X T4 DNA Ligase Buffer(NEB), 1µL T4 DNA Ligase(NEB), 50 ng vector DNA and 50 ng insert DNA were mixed. Water was added up to 20µL. The mixture was incubated at RT for 2h. DNA was purified again by QIAquick PCR Purification Kit.

Procedure for in vitro STxB labeling

STxB in PBS pH 7.4 (20 µL, 150µM) was added to a solution of NHS compound in certain equivalence(1:1, 1:3, 1:9, or 1:27). The mixture was incubated for 2h at RT and dialyzed against PBS at 4°C with 10 kDa MWCO dialysis cassettes(Pierce). The coupling was verified by Maldi-TOF.

Procedure for cell surface modification with NHS linked compound

STxB(1µM in culture medium) was bound to cells at 4°C for 45min(if necessary). After 3 times washing by PBS++, the cells were incubated with NHS linked compound in certain concentration in PBS++ at 4°C for 1h. Free compound was removed by 3 times washing by PBS++. Cells then followed different treatments as required.

DISCUSSION

1. Identification of cargo candidates of the retrograde route

In the course of this study, a list of retrograde cargo proteins could be identified using the SNAP-tag approach. These could be categorized into three groups: transporters, receptors, and kinases. Most of them have not been fully characterized in previous studies, especially concerning their intracellular trafficking behavior.

1.1 Integrin

Integrins are well characterized as the main proteins to mediate cell adhesion. They are transmembrane receptor, consisting of alpha and beta subunit, that bind extracellular matrix (ECM) proteins to link them to cytoskeletal proteins and signaling pathways inside the cells (Hynes 1992). Many studies have shown that integrins are involved in cell migration (Caswell and Norman 2006; Moissoglu and Schwartz 2006). Recent activity within this research field has illuminated that integrin trafficking contributes to cell migration in a regulated endocytic–exocytic trafficking mechanism. Integrins are internalized in clathrin-dependent or -independent pathway to early endosomes (Altankov and Grinnell 1995; Naslavsky, Weigert et al. 2003). Further, re-exocytosis of these integrins occours towards the leading edge of the migrating cell (Caswell and Norman 2008). In our study, alpha 3 integrin, beta 1 integrin, and others were found as hits in several proteomics experiments. Further studies are required to proof that these integrins can indeed use the retrograde transport route, and how this relates to their functions in cell adhesion and migration. Interestingly, in a study on TGN38, a well known retrograde transport protein that cycles between its steady state localization at the TGN and the plasma membrane, beta 1 integrin was found as an interacting partner, and a modification of the trafficking of TGN38 resulted in a change in the distribution of integrin as well (Wang and Howell 2000). The interaction between the retrograde protein TGN38 and integrin could be a molecular explanation of why integrin was identified in our

proteomics study. The retrograde transport of integrins may facilitate cell migration by polarized secretion.

1.2 Transferrin receptor

The discovery of transferrin receptor (TfR) among the hits was unexpected. TfRe binds iron loaded ferritin at the plasma membrane. After endocytosis, acidification leads to iron release. The TfR-ferritin complex is then recycled to the plasma membrane where apo-ferritin is released into the extracellular medium to start a new functional cycle (Grant and Donaldson 2009). TfR is often used in cell biology experiments as an archetypical marker of early endosomes. However, in the very early studies on TfR, researchers had initially focused on TfR trafficking to the Golgi. Anti-TfR antibody uptake assay showed that a small fraction of TfR was recycled to Golgi (Woods, Doriaux et al. 1986). Other early study on TfR showed that asialo-TfR could be resialylated after its deglycosylation at the plasma membrane, strongly suggesting that the TfR was transported from the plasma membrane to a sialyltransferase-containing compartment: the TGN (Snider and Rogers 1985). Notably, the half-time of resialylation of TfR was significantly lower than its transport route for iron uptake, which implicated a different pathway. TfR was one of the first proteins to be identified as a cargo protein of the retrograde transport route. The re-discovery of TfR as a retrograde cargo protein may prompt new studies into the functional implications and consequences of this alternative intracellular distribution pathway.

1.3 Ion transporters

Previous studies have shown that cells adapted to different extracellular ion or metabolite concentration by regulating the intracellular compartmentalization of

transporter proteins through retrograde transport. Well studied examples are the copper transporter Menkes (MNK protein) and the glucose transport Glut4 in mammalian cells, and the iron transporter Fet3p/Ftr1p in yeast (Petris and Mercer 1999; Strochlic, Setty et al. 2007). Interestingly, some ion transporter proteins like SLC39A14(Zip14) were found in our screening. SLC39A14 is a zinc transporter that was first described in 2005 (Taylor, Morgan et al. 2005). It is a member of the LZT (LIV-1 subfamily of ZIP zinc transporters) subfamily of zinc influx proteins. It was shown that the alteration of the function of zinc transporter proteins played a potential role in the progression of Alzheimer's disease (Lovell 2009). One study showed that overexpressing SLC39A14 increased the uptake of Zinc (Liuzzi, Aydemir et al. 2006). Apart from this, nothing is know on the regulation of this transporter and its intracellular compartmentalization. Considering the similarity in function with MNK and the fact that the protein came up as a hit in our screen, one may hypothesize that cells may regulated intracellular Zinc concentrations by controlling plasma membrane levels of SLC39A14 via retrograde transport.

1.4 Critical discussion

The protein candidates identified by SNAP-tag approach are supposed to be captured in TGN, because the SNAP-tag is fused with a TGN retained protein. However, we can not entirely exclude the possibility that GalT-GFP-SNAP “leaks” to other cell compartments like early endosomes. Notably, Retro-2 can inhibit STxB-BG capturing, indicating that at least for STxB-BG, the interaction with SNAP-tag occurs in TGN.

To confirm the hits are retrograde cargo proteins, biochemistry tools as well as IF analysis are still undergoing. Western blotting using antibody against candidate protein can confirm the capturing since the SNAP-tag is covalently linked to proteins and can cause shift of the band in gel. IF analysis can be used to trace the candidate.

The known retrograde transport markers like MPR and TGN38 are not found in the hits. One explanation is that the steady state plasma membrane levels of those proteins are too low. This is a limitation of sensitivity in our approaches especially for those with a low plasma membrane level.

In the comparative proteomics, we used the experimental proteome to compare with the control proteome that is not BG-PEG_9-NHS labeled. In this strategy, all the proteins detected in the control proteome are defined as contamination and eliminated, which might miss true hits. To obtain a quantative comparasin, we start to integrate SILAC to SNAP-tag approach.

2. STxB as a model to verify proteomics approaches

In the course of this study, STxB was used as a model protein. This had several advantages. Firstly, the mechanisms of STxB transport via the retrograde route are extensively studied, and many molecular tools such as inhibitors like small molecule Retro compounds and trafficking tags like sulfation-site have been developed in the course of these studies (Johannes, Tenza et al. 1997; Amessou, Popoff et al. 2006). This has strongly facilitated the development of different proteomics approaches. Secondly, due to its strong interaction with cells (nanomolar apparent dissociation constant), STxB could be considered as membrane protein once it has bound to its Gb3 receptor on the cell surface. This feature made it a perfect model to estimate the cell surface modification efficacy, which was a crucial step in the development of our proteomics approaches. Moreover, this estimation was quantitative. HeLa cells have $4*10^7$ STxB binding sites on on their surface (Falguieres, Mallard et al. 2001), which is high compared to endogenous membrane proteins like transferrin receptor (about $2*10^5$ binding sites per cell (Bridges and Smith 1985)). STxB was therefore a robust tool for the optimization of the proteomics approaches, before addressing the challenging task of looking for endogenous proteins.

As a bacterially expressed protein, STxB could be easily genetically modified, expressed and purified. The purified STxB provided us with the *in vitro* possibility to link chemical probes such as BG-NHS or bSuPeR to the protein as an initial step for the validation of these chemical tools and the underlying proteomics strategies. Another advantage of STxB compared with other retrograde transport proteins resides in its one-way transport pattern along the retrograde transport route. Endogenous cargos of the retrograde transport route often cycled between different organelles, while STxB strictly follows the retrograde route from the plasma membrane to the Golgi and the ER, via endosomes. Therefore, STxB was demonstrated to be a reliable tracer to verify the functionality of our probes and the validity of our experimental setups and strategies. Using STxB, we achieved proof-of-concept demonstrations of 4 different proteomics approaches.

3. Innovative proteomic approaches

3.1 Proteomics strategy for retrograde transport

The interest in retrograde transport has recently intensified as more proteins were found to participate in this pathway (Bonifacino and Rojas 2006). It seems likely that this is only the tip of the iceberg. If Wnt needs retrograde transport for its function, is this characteristics limited to this one morphogen? If Glut4, Menkes and Fet3p/Ftr1p need retrograde transport for their functions, is this characteristics limited to these transporters? If galectin-9 uses retrograde transport for its functional cycle, is this characteristics limited to this galectin, and could galectin-9 and possibly other galectin bring cargo proteins along? These and many other examples illustrate the fact that most likely, much still remains to be discovered.

Various methods and techniques have been developed to study the molecular mechanisms of retrograde transport (Johannes and Popoff 2008). However, tools for the systematic discovery for retrograde cargo proteins were lacking. Proteomics

appears as a competitive option in this context. Unlike biochemistry or cell biology techniques based on known proteins or complexes, proteomics can discover protein candidates in an unbiased “de novo” way which might give insight into unexpected cellular functions. Our study responds to this challenge through the development of functional proteomics approaches which were validated step-by-step.

Similar to other high through put strategies, large scale study in proteomics gives inspiring results with huge “noise”. In mass spectrometry based proteomics, cellular protein contamination during purification is often the main cause for the “noise” that is observed. In the SNAP-tag approach, we used a comparative strategy to eliminate some of the protein contaminations.

3.2 Comparison with traditional proteomics approaches

In the field of membrane biology, traditional proteomics approaches have been very successful in establishing the molecular identify of compartments and transport intermediates. A great challenge is now to take into consideration the fact that membrane proteomes are usually highly dynamic, with proteins trafficking from one compartment to another, signaling pathways that share common routes, proteins that are continuously recycled, and proteins that are simply present in multiple compartments at the same time. Besides, in order to analyze proteins in different compartment, strategies usually based on the enrichment of specific organelles (Brunet, Thibault et al. 2003; Huber, Pfaller et al. 2003), which is problematic since subcompartments often share similar physicochemical properties.

In our study, a “bait and trap” concept was introduced to extend the scope of traditional proteomics approaches. Plasma membrane proteins were covalently labeled by a chemical probe, and their trafficking was monitored by a capture step at the TGN. A specific purification step by immunoprecipitation was used instead of enriching to reduce contamination. Two aspects are keys for this model: proper

selection of the “bait” and “trap” pair, and precise localization of the “trap” in the target compartment. Within our interdisciplinary team, we could design and chemically synthesize the “bait” with desired characteristics, properly set the “trap” through the choice of appropriate genetic fusing proteins, follow trafficking of labeled proteins in cells, and finally screen for retrograde cargo candidates by mass spectrometry. This innovative concept of proteomics can in principle be applied to study other endocytosis transport events.

3.3 Comparison of the four approaches

Four approaches are developed in parallel using different “bait” and “trap” pairs. The SNAP-tag approach was eventually chosen to be pushed until the level of proteomics analysis. However, the other approaches also remain viable at this stage and may actually yield complementary information. Each of them has its own advantage and limitations.

The SNAP-tag approach is based on a chemical suicide reaction between the SNAP-tag and BG derivatives to form a covalent bound that is resistant even to strong detergent washes. This covalent bond facilitates purification. The FKBP and streptavidin approaches both are based on non-covalent interactions, and extensive washing may eventually cause disassociation. In sulfation strategy, a radioactive sulfate is covalently added onto the sulfation-site peptide, which like in the case of the SNAP-tag strategy provides a stable marker throughout purification.

One significant advantage of sulfation approach is that this strategy can be performed in any cell type without the need for generating stable cell lines. Indeed, tyrosylprotein sulfotransferases (TPSTs) exist in almost all animal species as well as in plants (Negishi, Pedersen et al. 2001; Moore 2009). This property should in principle allow for the use of the sulfation approach for proteomics in different cell types, such as normal versus cancer cells. Such comparative approaches would open

avenues for the understanding of a role for retrograde transport in complex cellular phenomena. Of interest, retrograde transport has been implicated in signal transduction by EGF receptors (Liao and Carpenter 2007). While the excat meaning of this finding remains unclear at this point, it opens the possibility that this or other signaling molecules with key functions in tumorigenesis could rely on retrograde trafficking for some of their functions. The other three approaches require fusion protein expression which is difficult to do on stem cells or primary cells taken directly from living tissue.

However, there are disadvantages of sulfation approach. It requires the usage of radioactivity which leads to more complicated experiments. To identify proteins, radioactive materials need to be injected to mass spectrometry machines, which might cause problems to MS platform.

Immunoprecipitation is necessary all the four approaches. In the sulfation approach, all cell surface bSuPeR-labeled proteins are pulled down. Proteins that take the retrograde transport route will be visualized by autoradiography. These will then have to be separated from all the others. In the other approaches, only the trap proteins and their interacting partners (covalent or not) are pulled down. Proper washing and comparative proteomics should allow to bring down background noise as much as possible.

The FKBP and SNAP-tag approaches are quite similar. The reason for keeping them going in parallel was the uncertainty notably on whether the SNAP-tag would react with full size BG-tagged bait proteins in the TGN/Golgi cisternae of living cells. One of the striking findings of this study is that this is actually the case. The SNAP-tag methodology then has the advantage that trap and bait are covalently linked, which should facilitate purification.

Unlike the other three approaches in which the trap is set in the Golgi, the streptavidin approach is designed to screen proteins taking the retrograde route to the ER, like Shiga toxin, ricin and Cholera toxin (Lord, Jolliffe et al. 2003; Bernasconi

and Molinari 2011). Few endogenous proteins have been found to be transported from the plasma membrane to ER, such as sulfatase modifying factor 1 (Zito, Buono et al. 2007). This project has been put on hold due to technical problems and the fact that efforts were concentrated on the SNAP-tag approach. Yet, the streptavidin approach should be followed up to focus on the ER question, and to extend the “bait and trap” concept to other transport pathways.

4. SNAP-tag for protein-protein interaction

SNAP-tag is a recent developed technique to label proteins using different fluorescent BG derivatives. For this, the SNAP-tag is fused onto target proteins (Damoiseaux, Keppler et al. 2001; Hinner and Johnsson 2010). SNAP-tag labeling has contributed to a number of recent studies, such as HIV-cell interactions (Eckhardt, Anders et al. 2011), intracellular movements of calcium ions (Palmer 2009), and histone protein distribution (Klein, Loschberger et al. 2011). Here, we have applied the SNAP-tag approach in a new context: the study of protein-protein interactions *in vivo*. We have shown that STxB-BG as well as surface labeled proteins couple to the SNAP-tag within the compartments of living cells. This is the first time that SNAP-tag is thus shown to catalyze the suicide reaction with a full size BG-labeled protein. This study thereby expands the SNAP-tag tool kit to the study of intracellular protein-protein interaction. The reaction has proven to be sufficiently sensitive, specific, and compatible with intra-organelle reaction conditions.

5. Chemical probe synthesis

Peptide-protein conjugates are wildly used in different fields of research (Temming, Meyer et al. 2007; Kumaresan, Luo et al. 2008). In our study, the sulfation-site peptide is introduced onto membrane proteins, thereby transforming

them into substrates for sulfotransferase. In collaboration with Dr. Melnyk, we have developed a peptide-protein conjugation method that is based on an original synthesis scheme to obtain peptide-NHS reagents (bSuPeR-carbamate, bSuPeR-ester and bSuPeR-R-ester). This is a one step reaction in which the NHS moiety is attached in a high yield via an unprotected lysine residue. These peptide-NHS conjugates are stable in standard RP-HPLC purification at low pH.

NHS can also be linked to peptides by the azide alkyne Huisgen cycloaddition method (click reaction). Compared with our lysine-based one-step NHS addition method, the azide alkyne Huisgen cycloaddition is more complicated to achieve. In a recent study, a new click reaction was reported for the attachment of small molecules to proteins or peptides under mild, aqueous reaction conditions. This reaction involves cyclic diazodicarboxamides that react efficiently and selectively with the phenolic side chain of tyrosine (Ban, Gavrilyuk et al. 2010). This method, which is as simple as our lysine-based scheme, cannot be used in our study due to the presence of a tyrosine in the sulfation-site peptide sequence.

6 Negative charge and polyarginine

The insufficient efficiency of surface modification with bSuPeR has been a stumbling block for the sulfation method, which historically was our first choice. The huge difference that we observed in terms of efficacy between bSuPeR-ester and bSuPeR-R-ester clearly points to the importance of charge in plasma membrane protein modification. This problem came up specifically with the sulfation approach as the sulfation-site peptide happens to be highly negatively charged, while in other three approaches the bait probes are hydrophilic, but neutral. We reported an efficient improvement for surface labeling by adding arginines to the sulfation peptide sequence.

PERSPECTIVE

In the course of this study, four different proteomics approaches to study retrograde transport are developed. Step-by-step verification is performed using multiple assays. In SNAP-tag approach, a list of proteins candidates is identified. Further experiments are required to confirm these proteins discovered by proteomics approaches. Thereby, successful development of proteomics approaches proposed in this study is the beginning rather than the end. Once, by biochemical and image analysis, the protein candidates are confirmed to take transport route from plasma membrane to Golgi apparatus, a new project will be raised for each of those proteins concerning its expression, distribution, signaling, function and regulation. The knowledge obtained from these studies in newly found retrograde cargo proteins can essentially contribute to understand the overall function of retrograde transport.

More efforts are required to improve the sensitivity of these approaches. As indicated in the introduction of tools to study retrograde transport, in some cases, the amount of the endogenous cargos is very few in cell surface which decreases the sensitivity of approaches using NHS group in cell surface modification, since the cargos are available for probe labeling only when they are present in cell surface. We expect more proteins to be discovered with sensitivity increased by multiple times of labeling or simply scaling up the experiment materials including cell and probe.

Later study can be focused on comparing retrograde transport proteomes between different cell lines (epithelia, neuroblasts, fibroblasts, lymphocytes…) to enrich our knowledge of the physiological role of this pathway. Moreover, proteomes comparison can be performed between normal and “abnormal” cells. For example, the proteome of cancer cells can be used to compare with that of normal cells to discover tumor biomarkers for diagnostics, therapies and even targeted drug delivering [172].

In collaboration with Dr.Melnyk, bSuPeR is successfully synthesized. We also uncover the charge effect in surface modification. To improve sulfation approaches sensitivity, the peptide in bSuPeR for sulfation can be modified. For example, tandem sulfation-site peptide can be inserted to bSuPeR to increase the number of sulfation sites resulting stronger signal in autoradiography. In addition, peptide can be

synthesized in branch structure as well to increase the number of sulfation site.

The streptavidin approach has not been well developed. In order to study trafficking pathway from plasma membrane to ER, SNAP-tag approach can be adapted by fuse Iip33 to SNAP-tag resulting ER resident SNAP-tag, followed similar step-by-step validation.

ACKNOWLEDGEMENTS

During my PhD, I am glad to have the chance to collaborate and exchange ideas with lots of colleagues. I really appreciate their help, suggestion, collaboration and their willingness to share their knowledge. I am especially thankful to the following persons:

My PhD supervisor Dr. Ludger Johannes for giving me the chance to pursue my PhD in his group, his advice and help to perform this study.

Dr. Michel Azoulay and Dr. Jean-Claude FLORENT for their guide and help in chemistry synthesis.

The jury members Prof. Patrick Charnay, Dr.Sandrine Sagan, Dr.Oleg Melnyk, Prof.Mike Lord, Dr.Mariano Casado to review and evaluate my work.

Dr. Sandrine Sagan and Dr. Thierry Galli for their suggestions as my tutors.

Romain Christiano for helping me start the sulfation approach.

Reda Mhidia and Dr.Oleg Melnyk for their helpful collaboration in bSuPeR synthesis.

Wolfgang Faigle and Florent Dingli for their contribution to mass spectrometry analysis.

Severine Divoux and Franck Perez for their technical and material support in my projects.

Kai Johnsson for his suggestions in SNAP-tag approach.

The Institut Curie, UMR 144/CNRS, FRM for their (financial and technical) support during my PhD.

All members of the “Trafic, signalisation et ciblage intracellulaires” group for their support during my entire PhD.

REFERENCES

Abraham, R. T. and G. J. Wiederrecht (1996). "Immunopharmacology of rapamycin." Annu Rev Immunol **14**: 483-510.

Aebersold, R. and M. Mann (2003). "Mass spectrometry-based proteomics." Nature **422**(6928): 198-207.

Ahle, S. and E. Ungewickell (1989). "Identification of a clathrin binding subunit in the HA2 adaptor protein complex." J Biol Chem **264**(33): 20089-20093.

Altankov, G. and F. Grinnell (1995). "Fibronectin receptor internalization and AP-2 complex reorganization in potassium-depleted fibroblasts." Exp Cell Res **216**(2): 299-309.

Amessou, M., V. Popoff, et al. (2006). "Measuring retrograde transport to the trans-Golgi network." Curr Protoc Cell Biol **Chapter 15**: Unit 15 10.

Arighi, C. N., L. M. Hartnell, et al. (2004). "Role of the mammalian retromer in sorting of the cation-independent mannose 6-phosphate receptor." J Cell Biol **165**(1): 123-133.

Baenziger, J. U. and D. Fiete (1979). "Structural determinants of Ricinus communis agglutinin and toxin specificity for oligosaccharides." J Biol Chem **254**(19): 9795-9799.

Ban, H., J. Gavrilyuk, et al. (2010). "Tyrosine bioconjugation through aqueous ene-type reactions: a click-like reaction for tyrosine." J Am Chem Soc **132**(5): 1523-1525.

Barlowe, C., L. Orci, et al. (1994). "COPII: a membrane coat formed by Sec proteins that drive vesicle budding from the endoplasmic reticulum." Cell **77**(6): 895-907.

Benedum, U. M., A. Lamouroux, et al. (1987). "The primary structure of human secretogranin I (chromogranin B): comparison with chromogranin A reveals homologous terminal domains and a large intervening variable region." EMBO J **6**(5): 1203-1211.

Bernasconi, R. and M. Molinari (2011). "ERAD and ERAD tuning: disposal of cargo and of ERAD regulators from the mammalian ER." Curr Opin Cell Biol **23**(2): 176-183.

Bettelheim, F. (1954). "Tyrosine-O-sulfate in a peptide from fibrinogen." J Am Chem Soc **76**(10): 2838–2839.

Bierer, B. E., P. S. Mattila, et al. (1990). "Two distinct signal transmission pathways in T lymphocytes are inhibited by complexes formed between an immunophilin and either FK506 or rapamycin." Proc Natl Acad Sci U S A **87**(23): 9231-9235.

Blagoveshchenskaya, A. D., L. Thomas, et al. (2002). "HIV-1 Nef downregulates MHC-I by a PACS-1- and PI3K-regulated ARF6 endocytic pathway." Cell **111**(6): 853-866.

Blank, G. S. and F. M. Brodsky (1987). "Clathrin assembly involves a light chain-binding region." J Cell Biol **105**(5): 2011-2019.

Blondeau, F., B. Ritter, et al. (2004). "Tandem MS analysis of brain clathrin-coated vesicles reveals their critical involvement in synaptic vesicle recycling." Proc Natl Acad Sci U S A **101**(11): 3833-3838.

Blot, G., K. Janvier, et al. (2003). "Targeting of the human immunodeficiency virus type 1 envelope to the trans-Golgi network through binding to TIP47 is required for env incorporation into virions and infectivity." J Virol **77**(12): 6931-6945.

Bonfanti, L., A. A. Mironov, Jr., et al. (1998). "Procollagen traverses the Golgi stack without leaving the lumen of cisternae: evidence for cisternal maturation." Cell **95**(7): 993-1003.

Bonifacino, J. S. and J. H. Hurley (2008). "Retromer." Curr Opin Cell Biol **20**(4): 427-436.

Bonifacino, J. S. and R. Rojas (2006). "Retrograde transport from endosomes to the trans-Golgi network." Nat Rev Mol Cell Biol **7**(8): 568-579.

Borner, G. H., M. Harbour, et al. (2006). "Comparative proteomics of clathrin-coated vesicles." J Cell Biol **175**(4): 571-578.

Brac, T. (1984). "The charge distribution in the rough endoplasmic reticulum/Golgi complex transitional area investigated by microinjection of charged tracers." Tissue Cell **16**(6): 859-871.

Brass, A. L., D. M. Dykxhoorn, et al. (2008). "Identification of host proteins required for HIV infection through a functional genomic screen." Science **319**(5865): 921-926.

Bridges, K. R. and B. R. Smith (1985). "Discordance between transferrin receptor expression and susceptibility to lysis by natural killer cells." J Clin Invest **76**(3): 913-918.

Brown, E. J., M. W. Albers, et al. (1994). "A mammalian protein targeted by G1-arresting rapamycin-receptor complex." Nature **369**(6483): 756-758.

Brunet, S., P. Thibault, et al. (2003). "Organelle proteomics: looking at less to see more." Trends Cell Biol **13**(12): 629-638.

Burgoyne, R. D. and M. J. Clague (2003). "Calcium and calmodulin in membrane fusion." Biochim Biophys Acta **1641**(2-3): 137-143.

Campos, C., M. Kamiya, et al. (2011). "Labelling cell structures and tracking cell lineage in zebrafish using SNAP-tag." Dev Dyn **240**(4): 820-827.

Caswell, P. and J. Norman (2008). "Endocytic transport of integrins during cell migration and invasion." Trends Cell Biol **18**(6): 257-263.

Caswell, P. T. and J. C. Norman (2006). "Integrin trafficking and the control of cell migration." Traffic **7**(1): 14-21.

Chaudhry, A., S. R. Das, et al. (2007). "A two-pronged mechanism for HIV-1 Nef-mediated endocytosis of immune costimulatory molecules CD80 and CD86." Cell Host Microbe **1**(1): 37-49.

Clardy, J. (1995). "The chemistry of signal transduction." Proc Natl Acad Sci U S A **92**(1): 56-61.

Collins, B. M. (2008). "The structure and function of the retromer protein complex." Traffic **9**(11): 1811-1822.

Damoiseaux, R., A. Keppler, et al. (2001). "Synthesis and applications of chemical probes for human O6-alkylguanine-DNA alkyltransferase." Chembiochem **2**(4): 285-287.

Danial, N. N., C. F. Gramm, et al. (2003). "BAD and glucokinase reside in a mitochondrial complex that integrates glycolysis and apoptosis." Nature **424**(6951): 952-956.

Diamandis, E. P. and T. K. Christopoulos (1991). "The biotin-(strept)avidin system: principles and applications in biotechnology." Clin Chem **37**(5): 625-636.

Duncan, J. R. and S. Kornfeld (1988). "Intracellular movement of two mannose 6-phosphate receptors: return to the Golgi apparatus." J Cell Biol **106**(3): 617-628.

Eaton, S. (2008). "Retromer retrieves wntless." Dev Cell **14**(1): 4-6.

Eckhardt, M., M. Anders, et al. (2011). "A SNAP-Tagged Derivative of HIV-1-A Versatile Tool to Study Virus-Cell Interactions." PLoS One **6**(7): e22007.

Elia, G. (2008). "Biotinylation reagents for the study of cell surface proteins." Proteomics **8**(19): 4012-4024.

Falguieres, T., F. Mallard, et al. (2001). "Targeting of Shiga toxin B-subunit to retrograde transport route in association with detergent-resistant membranes." Mol Biol Cell **12**(8): 2453-2468.

Fields, S. and O. Song (1989). "A novel genetic system to detect protein-protein interactions." Nature **340**(6230): 245-246.

Frenzen, P. D., A. Drake, et al. (2005). "Economic cost of illness due to Escherichia coli O157 infections in the United States." J Food Prot **68**(12): 2623-2630.

Fruman, D. A., M. A. Wood, et al. (1995). "FK506 binding protein 12 mediates sensitivity to both FK506 and rapamycin in murine mast cells." Eur J Immunol **25**(2): 563-571.

Fujinaga, Y., A. A. Wolf, et al. (2003). "Gangliosides that associate with lipid rafts mediate transport of cholera and related toxins from the plasma membrane to endoplasmic reticulm." Mol Biol Cell **14**(12): 4783-4793.

Futter, C. E., A. Pearse, et al. (1996). "Multivesicular endosomes containing internalized EGF-EGF receptor complexes mature and then fuse directly with lysosomes." J Cell Biol **132**(6): 1011-1023.

Ghosh, P., N. M. Dahms, et al. (2003). "Mannose 6-phosphate receptors: new twists in the tale." Nat Rev Mol Cell Biol **4**(3): 202-212.

Girard, M., P. D. Allaire, et al. (2005). "Non-stoichiometric relationship between clathrin heavy and light chains revealed by quantitative comparative proteomics of clathrin-coated vesicles from brain and liver." Mol Cell Proteomics **4**(8): 1145-1154.

Girotti, M. and G. Banting (1996). "TGN38-green fluorescent protein hybrid proteins expressed in stably transfected eukaryotic cells provide a tool for the real-time, in vivo study of membrane traffic pathways and suggest a possible role for ratTGN38." J Cell Sci **109 (Pt 12)**: 2915-2926.

Golgi, C. (1989). "On the structure of nerve cells. 1898." J Microsc **155**(Pt 1): 3-7.

Grant, B. D. and J. G. Donaldson (2009). "Pathways and mechanisms of endocytic recycling." Nat Rev Mol Cell Biol **10**(9): 597-608.

Graves, H. C. (1988). "The effect of surface charge on non-specific binding of rabbit immunoglobulin G in solid-phase immunoassays." J Immunol Methods **111**(2): 157-166.

Griffith, J. P., J. L. Kim, et al. (1995). "X-ray structure of calcineurin inhibited by the immunophilin-immunosuppressant FKBP12-FK506 complex." Cell **82**(3): 507-522.

Griscelli, C., A. Durandy, et al. (1978). "A syndrome associating partial albinism and immunodeficiency." Am J Med **65**(4): 691-702.

Gronemeyer, T., C. Chidley, et al. (2006). "Directed evolution of O6-alkylguanine-DNA alkyltransferase for applications in protein labeling." Protein Eng Des Sel **19**(7): 309-316.

Haigler, H. T., J. A. McKanna, et al. (1979). "Rapid stimulation of pinocytosis in

human carcinoma cells A-431 by epidermal growth factor." J Cell Biol **83**(1): 82-90.

Harada, H., B. Becknell, et al. (1999). "Phosphorylation and inactivation of BAD by mitochondria-anchored protein kinase A." Mol Cell **3**(4): 413-422.

Harter, C. and F. Wieland (1996). "The secretory pathway: mechanisms of protein sorting and transport." Biochim Biophys Acta **1286**(2): 75-93.

Haughland, R. (1997). "Molecular probes." Handbook of fluorescent probes and research chemicals **679**.

Haugland, R. P. and W. W. You (2008). "Coupling of antibodies with biotin." Methods Mol Biol **418**: 13-24.

He, X., F. Li, et al. (2005). "GGA proteins mediate the recycling pathway of memapsin 2 (BACE)." J Biol Chem **280**(12): 11696-11703.

Hermanson, G. T. (2008). "Bioconjugate Techniques (Second Edition)." Academic Press

Hinner, M. J. and K. Johnsson (2010). "How to obtain labeled proteins and what to do with them." Curr Opin Biotechnol **21**(6): 766-776.

Hirst, J., S. E. Miller, et al. (2004). "EpsinR is an adaptor for the SNARE protein Vti1b." Mol Biol Cell **15**(12): 5593-5602.

Hirst, J., A. Motley, et al. (2003). "EpsinR: an ENTH domain-containing protein that interacts with AP-1." Mol Biol Cell **14**(2): 625-641.

Huber, L. A., K. Pfaller, et al. (2003). "Organelle proteomics: implications for subcellular fractionation in proteomics." Circ Res **92**(9): 962-968.

Huisgen, R. (1961). "Centenary Lecture - 1,3-Dipolar Cycloadditions." Proceedings of the Chemical Society of London **357**.

Hynes, R. O. (1992). "Integrins: versatility, modulation, and signaling in cell adhesion." Cell **69**(1): 11-25.

James, P. (1997). "Protein identification in the post-genome era: the rapid rise of proteomics." Q Rev Biophys **30**(4): 279-331.

Johannes, L. and V. Popoff (2008). "Tracing the retrograde route in protein trafficking." Cell **135**(7): 1175-1187.

Johannes, L., D. Tenza, et al. (1997). "Retrograde transport of KDEL-bearing B-fragment of Shiga toxin." J Biol Chem **272**(31): 19554-19561.

Johannes, L. and C. Wunder (2011). "Retrograde Transport: Two (or More) Roads

Diverged in an Endosomal Tree?" Traffic **12**(8): 956-962.

Kaiser, C. and S. Ferro-Novick (1998). "Transport from the endoplasmic reticulum to the Golgi." Curr Opin Cell Biol **10**(4): 477-482.

Kebarle, P. and U. H. Verkerk (2009). "Electrospray: from ions in solution to ions in the gas phase, what we know now." Mass Spectrom Rev **28**(6): 898-917.

Kehoe, J. W. and C. R. Bertozzi (2000). "Tyrosine sulfation: a modulator of extracellular protein-protein interactions." Chem Biol **7**(3): R57-61.

Keppler, A., S. Gendreizig, et al. (2003). "A general method for the covalent labeling of fusion proteins with small molecules in vivo." Nat Biotechnol **21**(1): 86-89.

Keppler, A., M. Kindermann, et al. (2004). "Labeling of fusion proteins of O6-alkylguanine-DNA alkyltransferase with small molecules in vivo and in vitro." Methods **32**(4): 437-444.

Khan, S. M., B. Franke-Fayard, et al. (2005). "Proteome analysis of separated male and female gametocytes reveals novel sex-specific Plasmodium biology." Cell **121**(5): 675-687.

Kino, T., H. Hatanaka, et al. (1987). "FK-506, a novel immunosuppressant isolated from a Streptomyces. I. Fermentation, isolation, and physico-chemical and biological characteristics." J Antibiot (Tokyo) **40**(9): 1249-1255.

Kirchhausen, T. (2000). "Clathrin." Annu Rev Biochem **69**: 699-727.

Klein, T., A. Loschberger, et al. (2011). "Live-cell dSTORM with SNAP-tag fusion proteins." Nat Methods **8**(1): 7-9.

Klumperman, J., A. Hille, et al. (1993). "Differences in the endosomal distributions of the two mannose 6-phosphate receptors." J Cell Biol **121**(5): 997-1010.

Koltin, Y., L. Faucette, et al. (1991). "Rapamycin sensitivity in Saccharomyces cerevisiae is mediated by a peptidyl-prolyl cis-trans isomerase related to human FK506-binding protein." Mol Cell Biol **11**(3): 1718-1723.

Kouranti, I., M. Sachse, et al. (2006). "Rab35 regulates an endocytic recycling pathway essential for the terminal steps of cytokinesis." Curr Biol **16**(17): 1719-1725.

Kumar, N., G. Cha, et al. (2004). "Molecular complexity of sexual development and gene regulation in Plasmodium falciparum." Int J Parasitol **34**(13-14): 1451-1458.

Kumaresan, P. R., J. Luo, et al. (2008). "Evaluation of ketone-oxime method for developing therapeutic on-demand cleavable immunoconjugates." Bioconjug Chem **19**(6): 1313-1318.

Leyh, T. S. (1993). "The physical biochemistry and molecular genetics of sulfate activation." Crit Rev Biochem Mol Biol **28**(6): 515-542.

Leyris, J. P., T. Roux, et al. (2011). "Homogeneous time-resolved fluorescence-based assay to screen for ligands targeting the growth hormone secretagogue receptor type 1a." Anal Biochem **408**(2): 253-262.

Liao, H. J. and G. Carpenter (2007). "Role of the Sec61 translocon in EGF receptor trafficking to the nucleus and gene expression." Mol Biol Cell **18**(3): 1064-1072.

Lin, S. X., W. G. Mallet, et al. (2004). "Endocytosed cation-independent mannose 6-phosphate receptor traffics via the endocytic recycling compartment en route to the trans-Golgi network and a subpopulation of late endosomes." Mol Biol Cell **15**(2): 721-733.

Lippincott-Schwartz, J., T. H. Roberts, et al. (2000). "Secretory protein trafficking and organelle dynamics in living cells." Annu Rev Cell Dev Biol **16**: 557-589.

Liuzzi, J. P., F. Aydemir, et al. (2006). "Zip14 (Slc39a14) mediates non-transferrin-bound iron uptake into cells." Proc Natl Acad Sci U S A **103**(37): 13612-13617.

Lopez-Verges, S., G. Camus, et al. (2006). "Tail-interacting protein TIP47 is a connector between Gag and Env and is required for Env incorporation into HIV-1 virions." Proc Natl Acad Sci U S A **103**(40): 14947-14952.

Lord, M. J., N. A. Jolliffe, et al. (2003). "Ricin. Mechanisms of cytotoxicity." Toxicol Rev **22**(1): 53-64.

Lovell, M. A. (2009). "A potential role for alterations of zinc and zinc transport proteins in the progression of Alzheimer's disease." J Alzheimers Dis **16**(3): 471-483.

Lucocq, J. M., E. G. Berger, et al. (1989). "Mitotic Golgi fragments in HeLa cells and their role in the reassembly pathway." J Cell Biol **109**(2): 463-474.

Lueking, A., M. Horn, et al. (1999). "Protein microarrays for gene expression and antibody screening." Anal Biochem **270**(1): 103-111.

Lundblad, R. L. and R. A. Bradshaw (1997). "Applications of site-specific chemical modification in the manufacture of biopharmaceuticals: I. An overview." Biotechnol Appl Biochem **26 (Pt 3)**: 143-151.

Luzio, J. P., B. Brake, et al. (1990). "Identification, sequencing and expression of an integral membrane protein of the trans-Golgi network (TGN38)." Biochem J **270**(1): 97-102.

Mallard, F., C. Antony, et al. (1998). "Direct pathway from early/recycling endosomes

to the Golgi apparatus revealed through the study of shiga toxin B-fragment transport." J Cell Biol **143**(4): 973-990.

Mallard, F., B. L. Tang, et al. (2002). "Early/recycling endosomes-to-TGN transport involves two SNARE complexes and a Rab6 isoform." J Cell Biol **156**(4): 653-664.

Marcusson, E. G., B. F. Horazdovsky, et al. (1994). "The sorting receptor for yeast vacuolar carboxypeptidase Y is encoded by the VPS10 gene." Cell **77**(4): 579-586.

Mari, M., M. V. Bujny, et al. (2008). "SNX1 defines an early endosomal recycling exit for sortilin and mannose 6-phosphate receptors." Traffic **9**(3): 380-393.

Maxfield, F. R. and T. E. McGraw (2004). "Endocytic recycling." Nat Rev Mol Cell Biol **5**(2): 121-132.

Mayor, S. and R. E. Pagano (2007). "Pathways of clathrin-independent endocytosis." Nat Rev Mol Cell Biol **8**(8): 603-612.

McKanna, J. A., H. T. Haigler, et al. (1979). "Hormone receptor topology and dynamics: morphological analysis using ferritin-labeled epidermal growth factor." Proc Natl Acad Sci U S A **76**(11): 5689-5693.

Moissoglu, K. and M. A. Schwartz (2006). "Integrin signalling in directed cell migration." Biol Cell **98**(9): 547-555.

Monigatti, F., B. Hekking, et al. (2006). "Protein sulfation analysis--A primer." Biochim Biophys Acta **1764**(12): 1904-1913.

Moore, K. L. (2009). "Protein tyrosine sulfation: a critical posttranslation modification in plants and animals." Proc Natl Acad Sci U S A **106**(35): 14741-14742.

Motley, A., N. A. Bright, et al. (2003). "Clathrin-mediated endocytosis in AP-2-depleted cells." J Cell Biol **162**(5): 909-918.

Munier-Lehmann, H., F. Mauxion, et al. (1996). "Function of the two mannose 6-phosphate receptors in lysosomal enzyme transport." Biochem Soc Trans **24**(1): 133-136.

Naslavsky, N., R. Weigert, et al. (2003). "Convergence of non-clathrin- and clathrin-derived endosomes involves Arf6 inactivation and changes in phosphoinositides." Mol Biol Cell **14**(2): 417-431.

Negishi, M., L. G. Pedersen, et al. (2001). "Structure and function of sulfotransferases." Arch Biochem Biophys **390**(2): 149-157.

Niehrs, C. and W. B. Huttner (1990). "Purification and characterization of tyrosylprotein sulfotransferase." EMBO J **9**(1): 35-42.

Niehrs, C., W. B. Huttner, et al. (1992). "In vivo expression and stoichiometric sulfation of the artificial protein sulfophilin, a polymer of tyrosine sulfation sites." J Biol Chem **267**(22): 15938-15942.

Nilsson, T., M. Pypaert, et al. (1993). "Overlapping distribution of two glycosyltransferases in the Golgi apparatus of HeLa cells." J Cell Biol **120**(1): 5-13.

Olsen, J. V., B. Blagoev, et al. (2006). "Global, in vivo, and site-specific phosphorylation dynamics in signaling networks." Cell **127**(3): 635-648.

Olsnes, S. and A. Pihl (1972). "Ricin - a potent inhibitor of protein synthesis." FEBS Lett **20**(3): 327-329.

Orci, L., M. Stamnes, et al. (1997). "Bidirectional transport by distinct populations of COPI-coated vesicles." Cell **90**(2): 335-349.

Ozeran, J. D., J. Westley, et al. (1996). "Identification and partial purification of PAPS translocase." Biochemistry **35**(12): 3695-3703.

Palade, G. (1975). "Intracellular aspects of the process of protein synthesis." Science **189**(4206): 867.

Palmer, A. E. (2009). "Expanding the repertoire of fluorescent calcium sensors." ACS Chem Biol **4**(3): 157-159.

Pandey, A. and M. Mann (2000). "Proteomics to study genes and genomes." Nature **405**(6788): 837-846.

Paterson, A. D., R. G. Parton, et al. (2003). "Characterization of E-cadherin endocytosis in isolated MCF-7 and chinese hamster ovary cells: the initial fate of unbound E-cadherin." J Biol Chem **278**(23): 21050-21057.

Patterson, G. H., K. Hirschberg, et al. (2008). "Transport through the Golgi apparatus by rapid partitioning within a two-phase membrane system." Cell **133**(6): 1055-1067.

Pearse, B. M. (1976). "Clathrin: a unique protein associated with intracellular transfer of membrane by coated vesicles." Proc Natl Acad Sci U S A **73**(4): 1255-1259.

Petris, M. J. and J. F. Mercer (1999). "The Menkes protein (ATP7A; MNK) cycles via the plasma membrane both in basal and elevated extracellular copper using a C-terminal di-leucine endocytic signal." Hum Mol Genet **8**(11): 2107-2115.

Plante, M., S. Claveau, et al. (2008). "Mucolipidosis II: a single causal mutation in the N-acetylglucosamine-1-phosphotransferase gene (GNPTAB) in a French Canadian founder population." Clin Genet **73**(3): 236-244.

Poincloux, R., F. Lizarraga, et al. (2009). "Matrix invasion by tumour cells: a focus on MT1-MMP trafficking to invadopodia." J Cell Sci **122**(Pt 17): 3015-3024.

Popoff, V., G. A. Mardones, et al. (2007). "The retromer complex and clathrin define an early endosomal retrograde exit site." J Cell Sci **120**(Pt 12): 2022-2031.

Rabouille, C., N. Hui, et al. (1995). "Mapping the distribution of Golgi enzymes involved in the construction of complex oligosaccharides." J Cell Sci **108 (Pt 4)**: 1617-1627.

Raiborg, C. and H. Stenmark (2009). "The ESCRT machinery in endosomal sorting of ubiquitylated membrane proteins." Nature **458**(7237): 445-452.

Riggs, B., W. Rothwell, et al. (2003). "Actin cytoskeleton remodeling during early Drosophila furrow formation requires recycling endosomal components Nuclear-fallout and Rab11." J Cell Biol **163**(1): 143-154.

Rogaeva, E., Y. Meng, et al. (2007). "The neuronal sortilin-related receptor SORL1 is genetically associated with Alzheimer disease." Nat Genet **39**(2): 168-177.

Roth, J. and E. G. Berger (1982). "Immunocytochemical localization of galactosyltransferase in HeLa cells: codistribution with thiamine pyrophosphatase in trans-Golgi cisternae." J Cell Biol **93**(1): 223-229.

Rowe, R. G. and S. J. Weiss (2008). "Breaching the basement membrane: who, when and how?" Trends Cell Biol **18**(11): 560-574.

Saint-Pol, A., B. Yelamos, et al. (2004). "Clathrin adaptor epsinR is required for retrograde sorting on early endosomal membranes." Dev Cell **6**(4): 525-538.

Sandvig, K., J. Bergan, et al. (2010). "Endocytosis and retrograde transport of Shiga toxin." Toxicon **56**(7): 1181-1185.

Sano, T., S. Vajda, et al. (1998). "Genetic engineering of streptavidin, a versatile affinity tag." J Chromatogr B Biomed Sci Appl **715**(1): 85-91.

Saraste, J. and E. Kuismanen (1992). "Pathways of protein sorting and membrane traffic between the rough endoplasmic reticulum and the Golgi complex." Semin Cell Biol **3**(5): 343-355.

Saxon, E. and C. R. Bertozzi (2000). "Cell surface engineering by a modified Staudinger reaction." Science **287**(5460): 2007-2010.

Scheurer, S. B., C. Roesli, et al. (2005). "A comparison of different biotinylation reagents, tryptic digestion procedures, and mass spectrometric techniques for 2-D peptide mapping of membrane proteins." Proteomics **5**(12): 3035-3039.

Schreiber, S. T. D. a. S. L. (1997). "Single-Step Synthesis of Cell-Permeable Protein Dimerizers That Activate Signal Transduction and Gene Expression." J Am Chem Soc **119**(22): 5106-5109.

Schutze, M. P., P. A. Peterson, et al. (1994). "An N-terminal double-arginine motif maintains type II membrane proteins in the endoplasmic reticulum." EMBO J **13**(7): 1696-1705.

Seaman, M. N., J. M. McCaffery, et al. (1998). "A membrane coat complex essential for endosome-to-Golgi retrograde transport in yeast." J Cell Biol **142**(3): 665-681.

Shewan, A. M., E. M. van Dam, et al. (2003). "GLUT4 recycles via a trans-Golgi network (TGN) subdomain enriched in Syntaxins 6 and 16 but not TGN38: involvement of an acidic targeting motif." Mol Biol Cell **14**(3): 973-986.

Small, S. A. (2008). "Retromer sorting: a pathogenic pathway in late-onset Alzheimer disease." Arch Neurol **65**(3): 323-328.

Smith, M. J., H. M. Carvalho, et al. (2006). "The 13C4 monoclonal antibody that neutralizes Shiga toxin Type 1 (Stx1) recognizes three regions on the Stx1 B subunit and prevents Stx1 from binding to its eukaryotic receptor globotriaosylceramide." Infect Immun **74**(12): 6992-6998.

Snider, M. D. and O. C. Rogers (1985). "Intracellular movement of cell surface receptors after endocytosis: resialylation of asialo-transferrin receptor in human erythroleukemia cells." J Cell Biol **100**(3): 826-834.

Soskic, V., M. Gorlach, et al. (1999). "Functional proteomics analysis of signal transduction pathways of the platelet-derived growth factor beta receptor." Biochemistry **38**(6): 1757-1764.

Sphyris, N., J. M. Lord, et al. (1995). "Mutational analysis of the Ricinus lectin B-chains. Galactose-binding ability of the 2 gamma subdomain of Ricinus communis agglutinin B-chain." J Biol Chem **270**(35): 20292-20297.

Spooner, R. A., D. C. Smith, et al. (2006). "Retrograde transport pathways utilised by viruses and protein toxins." Virol J **3**: 26.

Stechmann, B., S. K. Bai, et al. (2010). "Inhibition of retrograde transport protects mice from lethal ricin challenge." Cell **141**(2): 231-242.

Steffen, A., G. Le Dez, et al. (2008). "MT1-MMP-dependent invasion is regulated by TI-VAMP/VAMP7." Curr Biol **18**(12): 926-931.

Strochlic, T. I., T. G. Setty, et al. (2007). "Grd19/Snx3p functions as a cargo-specific adapter for retromer-dependent endocytic recycling." J Cell Biol **177**(1): 115-125.

Studier, F. W. and B. A. Moffatt (1986). "Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes." J Mol Biol **189**(1): 113-130.

Tai, G., L. Lu, et al. (2004). "Participation of the syntaxin 5/Ykt6/GS28/GS15

SNARE complex in transport from the early/recycling endosome to the trans-Golgi network." Mol Biol Cell **15**(9): 4011-4022.

Taylor, K. M., H. E. Morgan, et al. (2005). "Structure-function analysis of a novel member of the LIV-1 subfamily of zinc transporters, ZIP14." FEBS Lett **579**(2): 427-432.

Teasdale, R. D., D. Loci, et al. (2001). "A large family of endosome-localized proteins related to sorting nexin 1." Biochem J **358**(Pt 1): 7-16.

Temming, K., D. L. Meyer, et al. (2007). "Improved efficacy of alphavbeta3-targeted albumin conjugates by conjugation of a novel auristatin derivative." Mol Pharm **4**(5): 686-694.

Thinakaran, G. and E. H. Koo (2008). "Amyloid precursor protein trafficking, processing, and function." J Biol Chem **283**(44): 29615-29619.

Tran, T. H., Q. Zeng, et al. (2007). "VAMP4 cycles from the cell surface to the trans-Golgi network via sorting and recycling endosomes." J Cell Sci **120**(Pt 6): 1028-1041.

Traub, L. M. and S. Kornfeld (1997). "The trans-Golgi network: a late secretory sorting station." Curr Opin Cell Biol **9**(4): 527-533.

Ungewickell, E. and D. Branton (1981). "Assembly units of clathrin coats." Nature **289**(5796): 420-422.

van der Sluijs, P., M. Hull, et al. (1992). "The small GTP-binding protein rab4 controls an early sorting event on the endocytic pathway." Cell **70**(5): 729-740.

Vezina, C., A. Kudelski, et al. (1975). "Rapamycin (AY-22,989), a new antifungal antibiotic. I. Taxonomy of the producing streptomycete and isolation of the active principle." J Antibiot (Tokyo) **28**(10): 721-726.

Waguri, S., F. Dewitte, et al. (2003). "Visualization of TGN to endosome trafficking through fluorescently labeled MPR and AP-1 in living cells." Mol Biol Cell **14**(1): 142-155.

Wahle, T., K. Prager, et al. (2005). "GGA proteins regulate retrograde transport of BACE1 from endosomes to the trans-Golgi network." Mol Cell Neurosci **29**(3): 453-461.

Wang, J. and K. E. Howell (2000). "The luminal domain of TGN38 interacts with integrin beta 1 and is involved in its trafficking." Traffic **1**(9): 713-723.

Wong, S. H., S. H. Low, et al. (1992). "The 17-residue transmembrane domain of beta-galactoside alpha 2,6-sialyltransferase is sufficient for Golgi retention." J Cell

Biol **117**(2): 245-258.

Woods, J. W., M. Doriaux, et al. (1986). "Transferrin receptors recycle to cis and middle as well as trans Golgi cisternae in Ig-secreting myeloma cells." J Cell Biol **103**(1): 277-286.

Yang, P. T., M. J. Lorenowicz, et al. (2008). "Wnt signaling requires retromer-dependent recycling of MIG-14/Wntless in Wnt-producing cells." Dev Cell **14**(1): 140-147.

Zito, E., M. Buono, et al. (2007). "Sulfatase modifying factor 1 trafficking through the cells: from endoplasmic reticulum to the endoplasmic reticulum." EMBO J **26**(10): 2443-2453.

Zolla, L., G. Lupidi, et al. (1990). "Encapsulation of proteins into human erythrocytes: a kinetic investigation." Biochim Biophys Acta **1024**(1): 5-9.

ANNEX

ANNEX 1

CHEMISTRY-BASED PROTEIN MODIfiCATION STRATEGY FOR ENDOCYTIC PATHWAY ANALYSIS

Christiano R, Amessou M, Shi G, Azoulay M, Blanpain A, Drobecq H, Melnyk O, Florent JC, Johannes L.

Biol Cell. 2010 Mar 25;102(6):351-9.

ANNEX 2

SYNTHESIS OF PEPTIDE-PROTEIN CONJUGATES USING N-SUCCINIMIDYL CARBAMATE CHEMISTRY

Mhidia R, Vallin A, Ollivier N, Blanpain A, Shi G, Christiano R,
Johannes L, Melnyk O.

Bioconjug Chem. 2010 Feb 17;21(2):219-28.