Biophysical analysis of protein-protein and protein-small molecule interactions

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Barbara Albertoni

aus

Locarno (CH)

Bonn 2011

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät
der Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Gutachter:	Prof. Dr. Michael Famulok
2. Gutachter:	Prof. Dr. Günter Mayer
Tag der Promotion:	20. Oktober 2011
Erscheinungsjahr:	2011

Solubility is the solution

Abstract

The validation of small molecule inhibitors identified by high throughput screening requires a set of robust assays for interaction analysis. Here I describe the implementation of three methods: bioluminescence resonance energy transfer (BRET) for the analysis of protein/protein interaction in cells, surface plasmon resonance (SPR) for the measurement of binding kinetics *in vitro* and capture compound mass spectrometry (CCMS) for the determination of binding specificity in proteome.

All the methods described were tested on cytohesins, a family of guanine nucleotide exchange factors. The very recent discovery of their additional role in the regulation of receptor tyrosine kinases (RTKs) signalling and the availability of specific small molecule inhibitors (the Secins) made them an interesting target.

BRET was applied to the analysis of a possible binding of the cytohesin ARNO to the RTK EGF receptor (EGFR). Two strategies were devised: a direct, where the interaction of ARNO and EGFR was monitored, and an indirect, which followed the changes in the EGFR/EGFR interaction upon overexpression of ARNO.

Two SPR approaches were developed to analyse the interaction between ARNO and the EGFR on the one hand, and to determine the kinetic parameters of binding of ARNO to its inhibitor Secin16 on the other hand.

For CCMS, a photoreactive affinity based SecinH3 probe (SecinH3-TPD) was synthesised and its ability to specifically label ARNO was shown, although the labelling yield was limited by low solubility. A protocol for the enrichment, digestion and MS-analysis of the labelled proteins was established.

Contents

Abbreviations

α_0	analyte (A) initial concentration
β	surface concentration of ligands (L)
γ	surface concentration of the complexes (AL)
D	diffusion coefficient
k_d	dissociation rate constant
k_a	association rate constant
k_M	mass transport coefficient
M_r	Molecular weight
V_C	Column volume
Da	Damkhöler number
ABP	Activity based probe
ABPP	Activity based protein profiling
ACN	Acetonitrile
ARF	ADP-ribosylation factor
ARNO	ARF nucleotide-binding site opener, cytohesin-2
ARNO-E156K	GEF inactive ARNO
ARNO-GFP	ARNO C-terminally tagged with GFP
ARNO-Sec7	Sec7 domain of ARNO
BFA	Brefeldin A
bio-ΔPBR	Biotinylated ARNO-ΔPBR
bio-Sec7	Biotinylated ARNO-Sec7
BRET	Bioluminescence resonance energy transfer
BSA	Bovine serum albumin
CCMS	Capture compound mass spectrometry
CDK	Cyclin-dependent kinase 2
CID	Collision induced dissociation
Da	Dalton
Diglyme	1-methoxy-2-(2-methoxyethoxy)-ethane
DTT	Dithiothreitol
EDC	1-ethyl-3-(3-dimethylaminopropyl) carbodiimide
EDTA	Ethylenediaminetetraacetic acid
EGFR	Epidermal growth factor receptor
EGFR-GFP	EGFR C-terminally tagged with GFP
EGFR-ICD	Cytoplasmic domain of the EGFR
EGFR-Luc	EGFR C-terminally tagged with luciferase
EGFR-Stop	EGFR construct without fluorophore
EGFRwE4	Full length EGFR, including the exon 4
eq.	Equivalent
ESI	Electrospray ionization
EtOH	Ethanol
FA	Formic acid
FBS	Fetal bovine serum
Fc	Flow cells
Fig.	Figure
FRET	Fluorescence resonance energy transfer
GAP	GTPase-activating proteins
GDP	Guanosine diphosphate
GEF	Guanine nucleotide exchange factor
GFP	Green fluorescent protein
GST	Glutathione S-transferase

ABBREVIATIONS

GST-GFP	GST C-terminally tagged with GFP
GST-Luc	GST C-terminally tagged with luciferase
GTP	Guanosine-5'-triphosphate
HEK cells	Human embryonic kidney 293 cells
HER2-Luc	HER2 C-terminally tagged with luciferase
HER3-GFP	HER3 C-terminally tagged with GFP
HPLC	High pressure liquid chromatography
IAA	Iodoacetamide
IC_{50}	Half maximal inhibitory concentration
IFC	Integrated microfluidic cartridge
IGFR	Insulin-like growth factor receptor
ITC	Isothermal titration calorimetry
LC-MS/MS	Liquid chromatography tandem mass spectrometry
Luc	*Renilla* luciferase
m/z	Mass-to-charge ratio
MALDI	Matrix-assisted laser desorption/ionization
MAP	Microtubule-associated protein
MIG6	Mitogen-induced gene 6
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
NHS	N-Hydroxysuccinimide
NTA	Nitrilotriacetic acid
o.n.	Overnight
PAGE	Polyacrylamide gel electrophoresis
PAZ	PAZ domain of Argonaute1
pI	Isoelectric point
PIP_2	Phosphatidylinositol 4,5-bisphosphate
PKC	Protein kinase C
PMF	Peptide mass fingerprinting
PS	Phosphatidylserine
PTM	Post-translational modification
Ref.	Reference
RET	Resonance energy transfer
RP	Reversed phase
RT	Room temperature
RTK	Receptor tyrosine kinase
RU	Resonance Unit
SDS	Sodium dodecyl sulfate
SecinH3-TPD	SecinH3 probe, modified with trifluorophenyl diazirine and desthiobiotin
SEM	Standard error of the mean
SPR	Surface plasmon resonance
Tab.	Table
TFA	Trifluoroacetic acid
TPD	Trifluorophenyl diazirine
TPD-COOH	4-(1-Azi-2,2,2-trifluoroethyl)benzoic acid
Tris	2-Amino-2-hydroxymethyl-propane-1,3-diol
w/v	Weight over volume
w/w	Weight over weight

1

Introduction

Small G proteins, or GTPases, are switches which regulate protein activity and localisation by cycling between an inactive guanosine diphosphate (GDP)-bound and an active guanosine triphosphate (GTP)-bound conformation. In their GTP-bound state they interact with effector proteins to induce downstream signalling events (Fig. 1.1).

The GDP-GTP cycle is highly regulated by various classes of proteins. Guanine nucleotide exchange factors (GEFs), for example, induce the release of bound GDP, which is then replaced by GTP, more abundant in the cell, while GTPase-activating proteins (GAPs) provide essential catalytic groups for GTP hydrolysis[1].

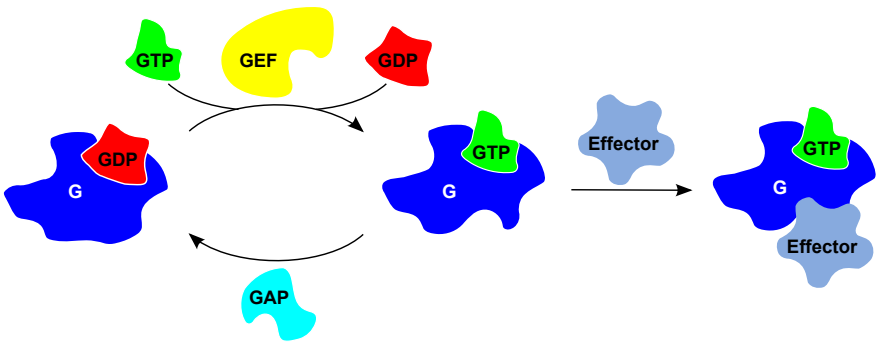

Figure 1.1: Regulation of GTPases - Small G proteins cycle between an inactive GDP-bound and an active GTP-bound state. Exchange of GDP for GTP is catalysed by guanine nucleotide exchange factors (GEFs), while GTPase-activating proteins (GAPs) induces hydrolysis of GTP back to GDP. In the active state, GTPases interact with effectors molecules to trigger downstream signalling.

GTPases are ubiquitously present and involved in various essential cellular processes, thus their disregulation is often cause of serious medical conditions. Indeed,

Ras, after which the Ras superfamily of small G proteins is named, was found to be aberrant in ∼33 % of the human tumours[2].

The more than 150 members of the Ras superfamily are divided in five major families and are key regulators of several aspects of normal cell growth and malignant transformation. While the members of the Ran family are responsible for nuclear import/export, nuclear envelope formation and spindle formation, Ras GTPases are mainly associated with transcription, cellular differentiation and proliferation. Rho family members control cell shape, cytoskeleton and cell migration, and Rab and Arf small G proteins are involved in vesicle-associated processes[1]. The adenosine diphosphate (ADP)-ribosylation factors (ARFs) family is of interest for this project and is presented in detail in next section.

1.1 ARF GTPases

ARFs were first described as a cellular activity required for cholera toxin to ADP-ribosylate an heterotrimeric G protein and exerts its toxic effect[3] but it was then readily identified as a GTP-binding protein[4] and its primary role in the regulation of vesicular transport is now well established[5].

The 6 members of the ARF family are divided in three classes. Class I comprises ARF1-3 and is involved in the assembly of vescicle coat complexes, ARF4-5 are part of Class II and the unique member of Class III, ARF6, regulates endocytosis, cytokinesis and cytoskeletal actin assembly[5,6]. While the class I and II ARFs are concentrated in the Golgi, ARF6 associates with the plasma membrane and a subset of endosomes at the cell periphery[5].

1.1.1 Structural features

ARFs have a mirystoylated N-terminal amphipathic helix, necessary for membrane binding, and two switch regions, which are responsible of binding of the effectors[7].

GDP-bound ARF is mainly cytosolic, while ARF-GTP is bound to the membrane. Indeed, exchange from GDP to GTP changes the conformation of the switch regions moving the interswitch away from the GTP-binding site and displacing the N-terminal helix from a hydrophobic pocket[7,8] (Fig. 1.2). This promotes insertion of the helix into the adjacent bilayer and assures that activation can exclusively occur at a membrane[7].

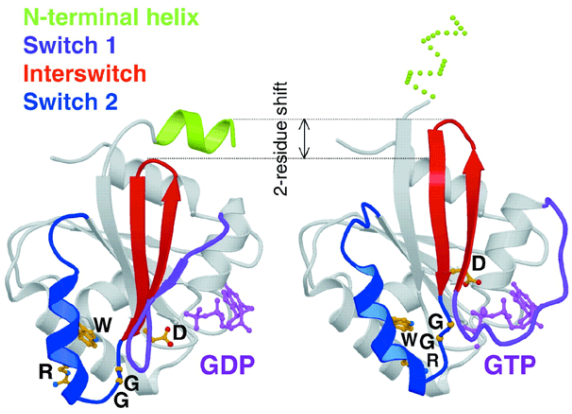

Figure 1.2: Structures of the ARF6 GDP- and GTP-bound states - Binding of GTP causes the interswitch (red) to move away from the switch regions (blue) The resulting displacement of the amphiphatic helix (green) favours its interaction with an adjacent lipid bilayer. *Reprinted by permission from Macmillan Publishers Ltd: EMBO Reports (Ref. 8), copyright 2002.*

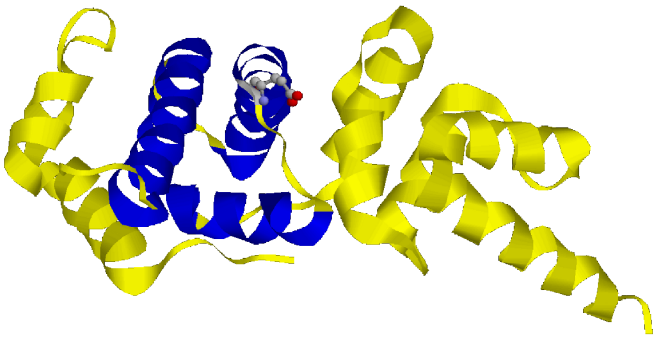

Figure 1.3: The Sec7 domain - Structure of the Sec7 domain of ARNO. The helices contributing to hydrophobic groove are shown in blue. The glutamic acid residue of the glutamic finger is shown in ball-and-stick representation. Image based on Protein Data Bank entry 1R8Q (Ref. 9).

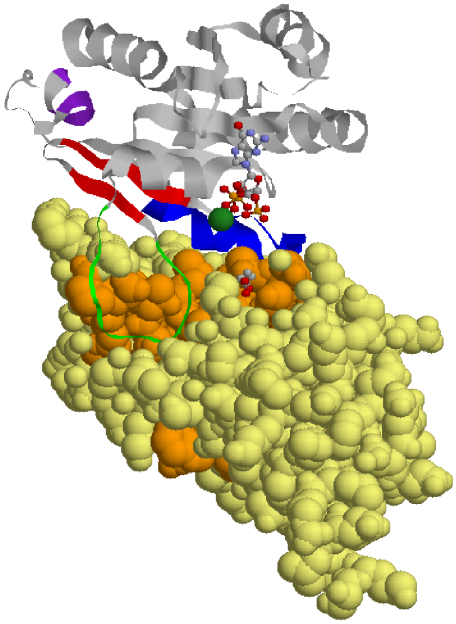

Figure 1.4: The ARF1-Sec7 complex - Structure of the complex of full length GDP-ARF1 (ribbons) and the Sec7 domain of a BFA-sensitive ARNO mutant (space-fill presentation) in the presence of BFA (not shown). GDP is shown in ball&sticks representation, Mg^{2+} in green. The conformational changes induced by ARNO binding on ARF, brings the glutamic finger (shown in ball&sticks) of ARNO in the nucleotide binding pocket of ARF, leading to displacement of the nucleotide. Since a BFA inhibited complex is shown, the glutamic finger does not reside in the binding pocket. Image based on Protein Data Bank entry 1R8Q (Ref. 9).

Table 1.1: ARF GEFs families and their localisation.[7]

Family	Class	Sec7 domain	Localisation
GBF1	large	yes	Golgi
BIG	large	yes	Golgi
PSD	small	yes	Plasma membrane
IQSEC	small	yes	Endosomes, cell periphery
Cytohesins	small	yes	Plasma membrane
FBX08	small	yes	
Sec12	small	no	ER, Golgi

1.1.2 ARF activation

Almost all ARF GEFs share a Sec7 domain with, in particular, a conserved hydrophobic groove and a glutamic finger which is necessary for catalysis[5–7] (Fig. 1.3). In the ARF/ARF GEF complex the switch 2 inserts in the hydrophobic groove while the switch 1 contacts the C-Terminus of the Sec7 domain. This induces substantial conformational changes in ARF that allow the glutamic finger to access the nucleotide-binding site and displaces the coordinating Mg^{2+} and the phosphate of the bound GDP[10–12]. This way, the glutamic finger extrudes the GDP from the binding pocket. The importance of the glutamic acid is confirmed by the fact that charge reversal at the glutamate residue reduces the exchange activity to background levels[11].

1.1.3 Guanine nucleotide exchange factors for ARFs

The ARF GEFs are divided in two major classes and 7 families based on sequence similarities and functional differences, as shown in Table 1.1.

1.1.3.1 GEF inhibitors

Due to their important regulatory functions, GEFs are interesting drug targets. However, up to now only few GEF inibitors are known[2]. The fungal toxin brefeldin A (Fig. 1.5), for example, inhibits ARF1 activation and therefore blocks many transport steps[13,14]. Brefeldin A (BFA) is an interfacial inhibitor which acts only on large GEFs and does not inhibit ARF6. It leads to loss of organelle identity, as for example the rapid disassembly of the Golgi complex and the redistribution of resident Golgi markers into the ER[15].

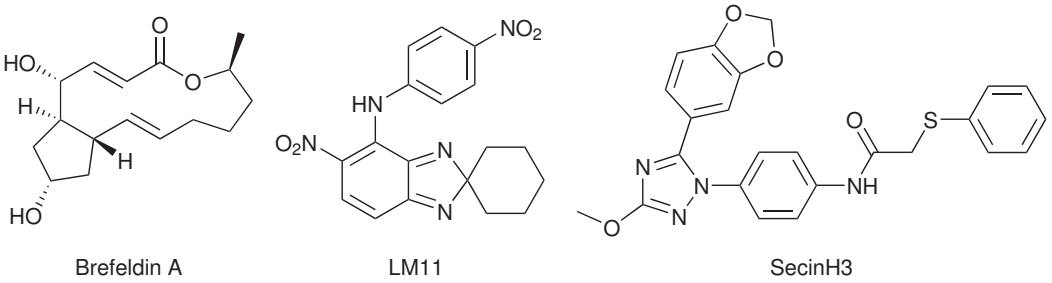

Figure 1.5: ARF GEF inhibitors - Structure of known small organic molecule ARF GEF inhibitors. Brefeldin A (Refs. 13 and 14) and LM11 (Ref. 16) are both interfacial inhibitors which bind the ARF/Sec7 complex. SecinH3 (Ref. 17) is a cytohesin specific inhibitor.

The second ARF GEF inhibitor, LM11 (Fig. 1.5), was identified only very recently by structure based screening, taking the BFA inhibition mechanism as model[16]. LM11 binds at the ARF1-Sec7 interface and inhibits both BFA-sensitive and -insensitives GEFs but does not inhibit ARF6.

1.2 Cytohesins

Up to very recently, no specific small GEF inhibitor was known. Thus, some years ago, our group took the BFA-insensitive cytohesin ARF GEF family as target for the development of new inhibitors which would allow the study of the functions of this protein class.

1.2.1 Structural features and functions

Cytohesin-1, cytohesin-2/ARNO*, cytohesin-3/Grp1 and cytohesin-4, the four known members of the cytohesin family, are highly homologous. The ∼47 kDa multidomain proteins have a coiled-coil (CC) amino terminal domain, which is thought to facilitate the interaction with cellular-binding partners, and a central Sec7 domain, which bears the GEF function[18]. The pleckstrin homology (PH) domain and the C-terminal region, which contains a large proportion of basic amino acids (polybasic region, PBR), are important for the regulation of association to the membrane by binding to inositol phospholipids[18] and were recently shown to autoinhibit the exchange activity of the

*ARNO: ARF nucleotide-binding site opener

Sec7-domain[19] (Fig. 1.6). In the polybasic region lye also protein kinase C (PKC) sites which are specifically phosphorylated in response to phorbol ester stimulation [20,21]

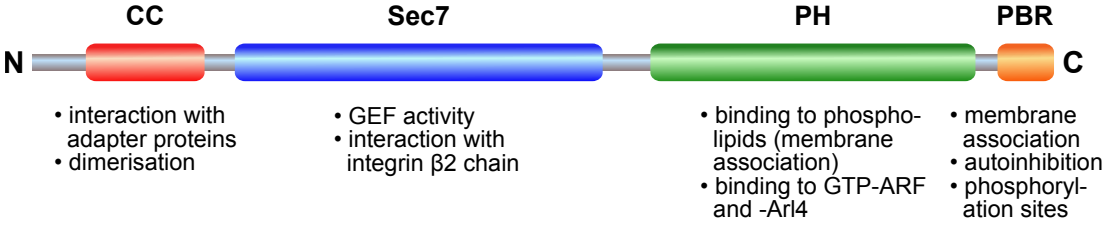

Figure 1.6: Cytohesins conserved domains and their function - CC: coiled-coil; PH: pleckstrin homology; PBR: polybasic region.

Additionally to their function as ARF GEFs, cytohesins were shown to interact with a variety of cell surface proteins, such as the β2 integrin LFA-1[22], the HHV8 protein kaposin A[23], the V-ATPase V0 complex[24] and the A_{2A} adenosine receptor[25], and are involved in MAPK signalling and in the phosphatidylinositol 3-kinase (PI3K) cascade[18].

1.2.1.1 Autoregulation of cytohesins

In the last years, a new level of regulation of exchange of cytohesins on ARF was discovered. In 2007, DiNitto *et al.* revealed the important role of the C-terminal polybasic region of cytohesins in their autoregulation[19]. Indeed, under resting conditions the docking site of the ARF switch regions on cytohesins is blocked by the linker region between Sec7 and PH-domain and by the C-terminal amphipathic helix. The resulting inhibition can be released by phosphoinositides-dependent binding of active ARF6 to the PH domain-PBR and makes thus activation of cytohesins dependent on plasma membrane recruitment[19] (Fig. 1.7).

The results above, together with the discovery that the cytohesin PH-domain binds the activated forms of ARF6 and Arl4 (Refs. 28–31), lead to the development of a new model for the mechanism of activation of ARF by ARNO[26]. According to this model, the requirement of binding of both anionic lipids (e.g. PS and PIP_2*) at the plasma membrane and ARF-GTP to release cytohesin autoinhibition impedes unlocalised activation. Then again, activation by ARF-GTP triggers a positive feedback cycle in

*PS: phosphatidylserine; PIP_2: phosphatidylinositol 4,5-bisphosphate

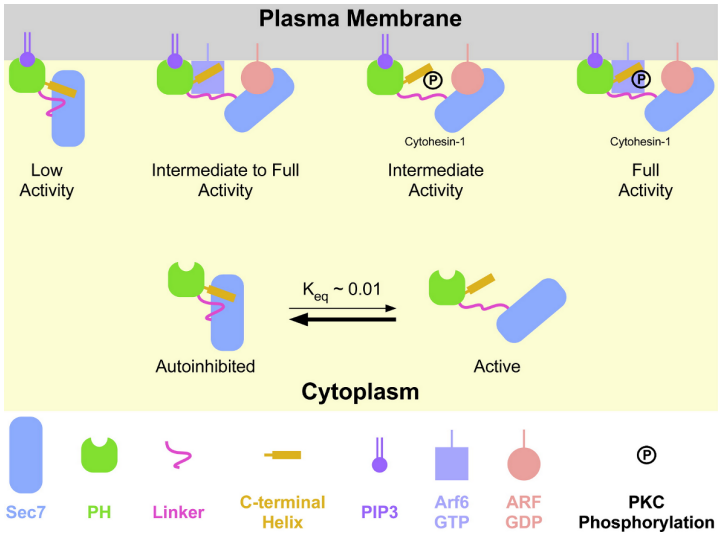

Figure 1.7: First model for autoregulation of cytohesins - After phosphoinositides-dependent plasma membrane recruitment of cytohesins, lateral association with ARF6-GTP simultaneously enhances membrane partitioning and shifts the equilibrium toward the catalytically competent conformation. Other mechanisms, including phosphorylation of PKC sites in the polybasic motif of cytohesin-1, may be required for full activation.[19]
Reprinted from Molecular Cell (Ref. 19), copyright 2007, with permission from Elsevier.

which ARNO activity is stimulated by its own product and allows a quick response to stimulating events[26] (Fig. 1.8).

1.2.2 The cytohesin inhibitor SecinH3

By mean of a newly developed aptamer displacement screening, which looked for small organic molecules capable of efficiently compete the specific binding of the aptamer M69 (Ref. 32) to the Sec7 domain of cytohesin-1, the cytohesin inhibitor SecinH3 (Fig. 1.5) was identified[17,33].

SecinH3 was shown to be selective for cytohesins and to inhibit the exchange of ARNO on ARF1 with an half maximal inhibitory concentration (IC_{50}) of ~11 μM (Refs. 17 and 34). More importantly, SecinH3 allowed to discover new roles of cytohesins.

1.2.3 Newly discovered roles of cytohesins

Application of SecinH3 in cell culture, mice and flies revealed the involvement of cytohesins in insulin signalling[17,35].

1. INTRODUCTION

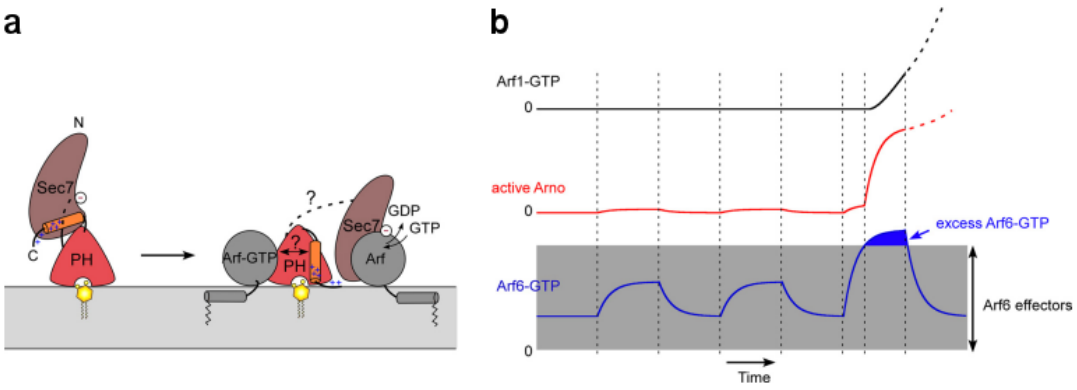

Figure 1.8: Structural and functional aspects of ARNO activation on lipid membranes - a. Proposed conformational changes in ARNO at the surface of lipid membranes. The switch to an active membrane-bound conformation depends not only on the presence of anionic lipids (e.g. PS + PIP_2) but also on the presence of a free ARF-GTP molecule. **b.** Model of ARNO response. In resting cells, ARNO remains inert because no membrane contains both anionic lipids and free ARF-GTP. Notably, the plasma membrane, which contains PS and PIP_2, might not be suitable for ARNO activation when active ARF species (e.g. ARF6) are engaged in interaction with classical effectors and for constitutive functions. Ignition of ARNO requires a burst of active ARF6 or Arl4 at the plasma membrane that exceeds the buffering capacity of effectors. After this initiation step, ARNO is engaged in a sustained self-activating pathway through the feedback effect of newly formed ARF-GTP molecules (e.g. ARF1) and remains active even if some initial inputs disappear. The proposed circuit is not restricted to ARF6 and ARF1 and could also apply when ARNO is activated by a different ARF subtype (e.g. Arl4).[26] *Reprinted from Journal of Biological Chemistry(Ref. 26), copyright 2011 The American Society for Biochemistry and Molecular Biology.*

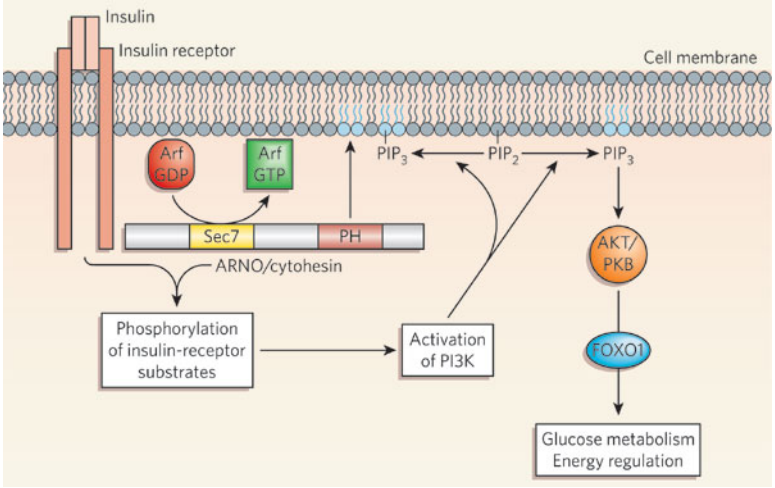

Figure 1.9: Cytohesins in the insulin-signalling cascade - Through formation of a complex with the insulin receptor and ARF6, cytohesins facilitate phosphorylation of IRS and the following signalling cascade. *Reprinted by permission from Macmillan Publishers Ltd: Nature (Ref. 27), copyright 2006.*

1.2.3.1 Regulation of insulin signalling

The insulin receptor is a receptor tyrosine kinase which upon binding of insulin undergoes autophosphorylation on cytoplasmic tyrosine residues and stimulates phosphorylation of the insulin receptor substrate proteins (IRS). These, in turn, stimulate the conversion of PIP_2 to PIP_3 (phosphatidylinositol 3,4,5-trisphosphate) by the phosphatidylinositol 3-kinase (PI(3)K) and lead to the activation of a number of downstream targets including FOXO1, a factor that regulates the expression of insulin-sensitive genes[17,27] (Fig. 1.9).

The effect of SecinH3 in cell culture, mice and flies was very similar to that obtained by impairment of the insulin receptor and it was shown that, although cytohesins do not affect insulin receptor phosphorylation, they still act very far upstream in the signalling pathway, probably by facilitating the binding of IRS to the insulin receptor[17,35]. This results showed for the first time that cytohesins play an essential role in insulin signalling and described a new ARF-dependent role of cytohesins.

1.2.3.2 Involvement in EGFR signalling

At the time when this project was started, preliminary results indicated that cytohesins are involved also in the regulation of ErbB receptor signalling. In her Diploma thesis,

A. Bill showed that SecinH3 treatment reduces the stimulation dependent activation of the ErbB receptor in cell culture and SecinH3-dependent inhibition was also observed in the downstream signalling cascade[36]. Moreover, inhibition of cytohesins led to a reduction of the proliferation of the lung cancer cell line H460[36]. This results suggested an important role of cytohesins in receptor tyrosine kinase signalling.

1.3 Receptor tyrosine kinases - The ErbB receptor family

Cell surface receptors regulate the proliferation of cells by recognising extracellular growth factors and transducing the signals across the cell membrane[37]. Receptor tyrosine kinases (RTKs) are a main class of cell surface receptors that catalyse the transfer of the γ phosphate of ATP to the hydroxyl groups of tyrosine residues on target proteins. Inactive RTKs exist as monomers in the cell membrane. They are activated by ligand-induced dimerisation resulting in autophosphorylation of their cytoplasmic domains, which then can act as a platform for the activation of signalling molecules further downstream[38]. RTKs are important regulators of fundamental cellular processes like metabolism, proliferation and differentiation. Aberrant signalling of the normally tightly controlled RTKs results in deregulated activity of downstream kinases and is a common feature of malignant transformations[39].

1.3.1 ErbB receptors dimerisation and activation

The members of the epidermal growth factor receptor family (EGFR or ErbB family) are typical RTKs with an extracellular ligand binding domain connected by a single transmembrane helix to the cytoplasmic domain[38]. The latter contains the protein tyrosine kinase (PTK) core and the carboxy-terminal tail with tyrosine autophosphorylation sites[40]. The phosphorylated residues serve as docking sites for signalling molecules that are responsible for the downstream transmission of the signal[41].

Upon binding to a ligand, the four ErbB family members EGFR/ErbB1/HER, ErbB2/HER2, Erb3/HER3 and ErbB4/HER4 can form homo- and heterodimers with any other ErbB receptor[40]. The formation of heterodimers is especially essential for ErbB2 and ErbB3, because, while no activating ligands are known for ErbB2, ErbB3 lacks intrinsic kinase activity. Therefore, both receptors are unable to support signalling as homodimers. Together, however, they form the most potent dimer in this receptor

class[40]. The potency of the transduced signal depends on the cellular context, the specific ligand, and the ErbB dimer[40].

However, it was reported that dimerisation can occur also in absence of ligand[42]. Thus, an additional level of regulation is necessary to avoid ligand independent activation of EGFR signalling. In fact, dimerisation alone is not sufficient to activate the ErbB receptors. The EGFR kinase domain is normally found in an autoinhibited conformation, similar to that first seen in the cyclin-dependent kinase 2 (CDK) and the Src family kinases[41,43]. Activation of the kinase requires the formation of an asymmetric dimer between the bottom of the C-lobe of one kinase monomer, which acts as a cyclin-like allosteric activator, and the top of the N-lobe of the other[41,43] (Fig. 1.10).

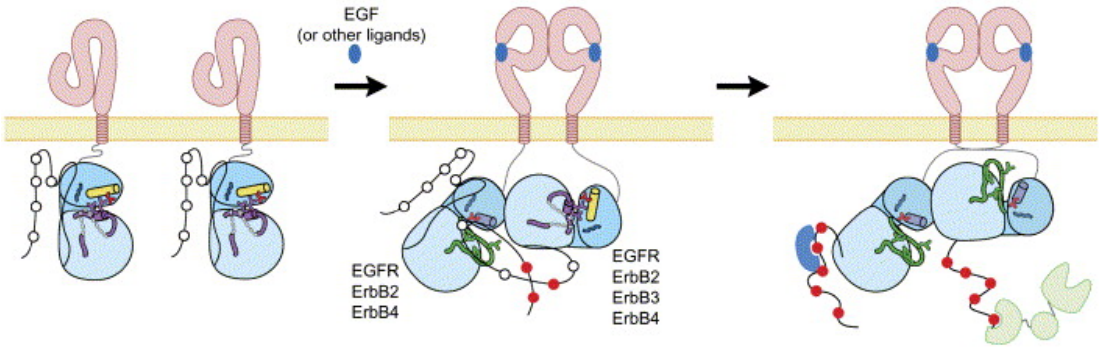

Figure 1.10: Model of the activation mechanism of the ErbB receptors - All the members in the family can act as the cyclin-like activator for the kinase-active members (EGFR, ErbB2, and ErbB4) after ligand-induced homo- or heterodimerization.[41] *Reprinted from Cell (Ref. 41), copyright 2006, with permission from Elsevier.*

Downregulation of EGFR signaling is regulated by internalization, dephosphorylation and degradation of the activated receptors[43]. Additionally, the mechanism of action of MIG6 (mitogen-induced gene 6), a feedback inhibitor which binds directly to the cytoplasmic region of the ErbB receptors, was recently described. MIG6 binds to the bottom of the kinase C-lobe, far away from the active site and blocks the EGFR kinase domain in the CDK/Src-like inactive conformation, preventing the formation of the active asymmetric dimer[43,44].

2

Aims of the project

The validation of small molecule inhibitors identified by high throughput screening requires robust assays for interaction analysis. The goal of this project was to establish a set of suitable assays to investigate both protein/small molecule and protein/protein interactions.

Cytohesins have been found to have a dual function as guanine nucleotide exchange factors (GEFs) on one side and ErbB receptor activators on the other side. Accordingly, the small molecule inhibitor SecinH3, which was developed as a GEF inhibitor acting on the cytohesins Sec7 domain, proved to be able to inhibit ErbB receptors' activation as well. This interesting characteristic offered the opportunity to develop both types of assays on the same target protein and at the same time to elucidate these distinct activities of cytohesins and better understand the mechanism of action of their inhibitors.

Three methods were chosen for implementation: bioluminescence resonance energy transfer (BRET) for the analysis of protein/protein interaction in cells, surface plasmon resonance (SPR) for the measurement of binding kinetics *in vitro* and capture compound mass spectrometry (CCMS) for the determination of binding specificity in proteome.

3

Bioluminescence resonance energy transfer

Preliminary results showing inhibition of ErbB receptor signalling by SecinH3[36] suggested cytohesins interact with ErbB receptors during their activation. To analyse this interaction and its possible function for receptor activation, a method for the analysis of protein-protein interaction was required. The use of Bioluminescence Resonance Energy Transfer (BRET) to study the interaction between RTKs and their downstream effector proteins in living cells had recently been reported by Tan *et al.*[37] and seemed to be an appropriate method for our purposes.

3.1 Principle of BRET

BRET, like FRET (Fluorescence Resonance Energy Transfer), involves the nonradiative (dipole-dipole) transfer of energy from a donor (an enzyme or fluorophore) to a complementary acceptor fluorophore, which usually emits fluorescence of a different wave length[45,46]. The efficiency of the energy transfer depends on the spectral overlap, the relative orientation of the dipoles, and the distance between the donor and acceptor fluorophores[45]. The efficiency of energy transfer is inversely proportional to the sixth power of this distance and RET effectively occurs when the donor and acceptor are 10-100 Å apart[45,47]. As this distance is comparable to the dimensions of macromolecular protein complexes and proteins can easily be expressed with a fluorescent tag, RET is a suitable method to analyse protein-protein interactions[46].

In BRET, the donor fluorophore of the FRET technique is replaced by a luciferase[45], an enzyme that oxidizes a substrate, such as coelenterazine, to emit light. Thus, BRET, not needing external excitation, avoids some of the problems associated with FRET, such as photobleaching, autofluorescence and simultaneous excitation of both donor and acceptor fluorophores[45,46].

Dependent on the substrate used, the luciferase reaction has emission peaks between 395 and 475 nm (Ref. 46), and the acceptor fluorophore, usually a GFP (green fluorescent protein) variant, has to be chosen accordingly. For this project a BRET^2 combination consisting of *Renilla* luciferase (RLuc) with the coelenterazine derivative DeepBlueC substrate as donor and GFP^2 as acceptor fluorophore[48] was used. The advantage of this combination is the superior separation of donor and acceptor emission peaks and thus a reduction of the background signal[46,49] (Fig. 3.1), while the drawbacks are the low quantum yield and rapid decay of DeepBlueC, with the resulting need of more cells and highly sensitive instruments to achieve sufficiently high luminescence levels for BRET detection[49].

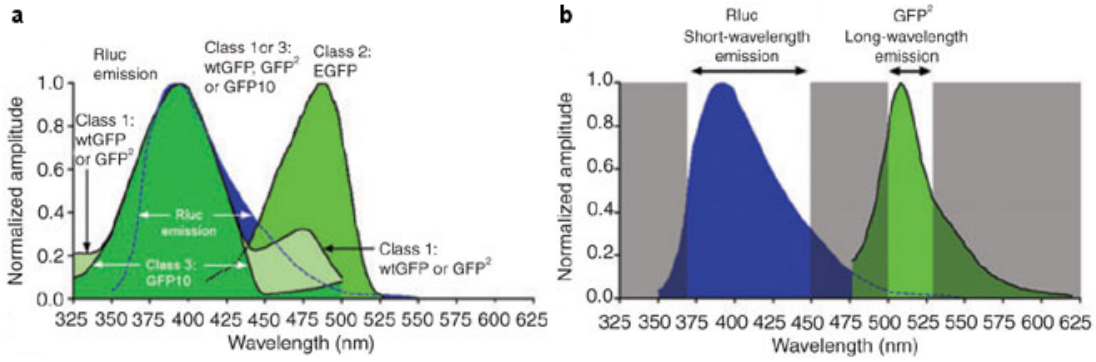

Figure 3.1: Schematic representation illustrating the overlap of normalized Rluc emission spectrum with normalized GFP excitation and emission spectra, together with a typical filter combination. - **a.** Overlap of Rluc emission spectrum with Class 1-3 GFP excitation spectra when using DeepBlueC. **b.** Lack of overlap of Rluc emission spectrum with GFP^2 emission spectra when using DeepBlueC. A typical filter set for the BRET^2 combination is shown. *Adapted by permission from Macmillan Publishers Ltd: Nature Methods (Ref. 46), copyright 2006.*

To investigate protein-protein interactions by BRET, it is necessary to express the proteins of interest fused to RLuc and GFP^2, respectively. For BRET analysis in living cells, cell populations coexpressing both fusion proteins are needed. To correct for RLuc emission at GFP wavelength, a control cell population expressing the donor-tagged

protein only is used. For both samples, light emission at both the short- (RLuc em.) and long-wavelength (GFP em.) is measured directly after addition of the substrate. The BRET signal, or BRET-ratio, is defined as the acceptor emission relative to the donor emission subtracted by the light detected at long-wavelength which is contributed by the donor emission (em.)[46,49]:

$$\text{BRET-ratio} = \frac{\text{GFP em.}}{\text{RLuc em.}} - \frac{\text{GFP em. of donor-only transfected cells}}{\text{RLuc em. of donor-only transfected cells}}. \tag{3.1}$$

This way, a high BRET-ratio generally indicates interaction between the two proteins. Analysis of the data with this formula controls also for variation in substrate and sample concentration[46].

There is no absolute correlation between BRET-ratio and strength of binding, thus the BRET signal of different proteins pairs can not be compared and can only be ascribed to binding by adequate control experiments, as described in the next section.

3.2 Results and discussion

The BRET assay offers two complementary approaches for the investigation of the ARNO-receptor interaction: binding of ARNO to the receptor can be measured directly, using luciferase tagged receptor and GFP labelled ARNO, or indirectly, by coexpressing non-labelled ARNO with luciferase and GFP tagged receptors. Both strategies were used in this project as described below.

3.2.1 Direct interaction analysis

To investigate the EGFR-ARNO interaction directly, two fusion proteins shall be expressed. Since ARNO can bind the EGFR only in its cytoplasmic region, the EGFR was constructed to carry *Renilla* luciferase (RLuc) at its C-terminus (EGFR-Luc), while ARNO was C-terminally fused to GFP (ARNO-GFP). The GEF inactive ARNO mutant ARNO-E156K[11], was used to verify if the interaction EGFR-ARNO is dependent on ARNO GEF activity.

As control, two proteins known to interact were needed. Thus, expression vectors for glutathione S-transferase (GST), a constitutively dimerised protein, fused to either RLuc (GST-Luc) or GFP (GST-GFP) were constructed. The same constructs can be

used to check for unspecific interaction when coexpressed with EGFR-Luc or ARNO-GFP.

The fusion proteins were transiently expressed by transfection of HEK cells. DNA of the GFP-tagged proteins was used in 60-times excess since this ratio showed the best results in preliminary exeriments. To avoid interference with the fluorescence readout, growing medium without phenol red was used. After ∼36 hours, the cells were harvested and transferred in a microplate. The plate reader was set up to inject the luciferase substrate (DeepBlueC) immediately before measurement. The BRET-ratio was calculated as described in Equation 3.1, p. 17.

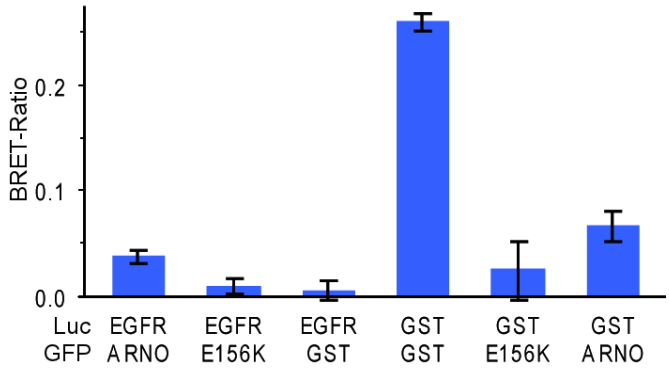

Figure 3.2: ARNO-EGFR interaction and GST controls - EGFR-Luc or GST-Luc were transiently coexpressed with either ARNO-GFP, ARNO-E156K-GFP or GST-GFP in excess in HEK cells. The BRET-ratio is given. The GST-Luc/GFP pair gave, as expected, a strong signal. While the EGFR-Luc/GST-GFP control showed no signal, a relatively high BRET-ratio was detected when GST-Luc was coexpressed with the ARNO mutants. Since the signal for the investigated pair (EGFR-Luc/ARNO-GFP) was lower than that of the negative control, it could not be attributed to specific interaction. n = 5 (E156K: n = 3).

As shown in Figure 3.2, the positive control cells, transfected with both GST constructs gave a strong BRET signal, supporting the experimental procedure used. However, the BRET-ratio obtained for the EGFR-ARNO pair was much lower. In fact, the negative control sample, where GST-Luc and ARNO-GFP were coexpressed, resulted in higher signals. This result speaks against a detection of specific binding in the EGFR-ARNO pair.

A second possibility to check specificity in a BRET system is to perform a competition assay where the RLuc and GFP tagged proteins are coexpressed with unlabelled protein[46]. This competes for interaction with the tagged protein, reducing the signal.

Trying to carry out this experiment with untagged ARNO as competitor, revealed one of the biggest problems of the experimental design. Indeed, the fluorescence signals detected were generally quite near to the background. Since calculation of the BRET-ratio involves quotients of these signals, a low signal to noise ratio gravely affects the results. This problem was particularly evident in the competitions experiments, since the amount of fluorescent protein was limited by the maximal quantity of DNA accepted by the cells. For the same reason, it was not possible to add the competing protein in excess, making competition per se more difficult. Hence, the results of the competition experiments were quite variable and difficult to interpret. Concurrently, the ARNO-GFP construct was not functional (A. Bill, J. Theis, unpublished results), therefore these experiments were suspended in favour of the indirect method.

3.2.2 Indirect interaction analysis

Monitoring the effects of ARNO on the BRET signal of the RLuc/GFP receptor pair offers the opportunity to observe the ARNO-EGFR interaction indirectly. For this strategy a second EGFR construct, carrying GFP at the C-terminus (EGFR-GFP), was needed. Additionally a second BRET pair with HER2 as donor (HER2-Luc) and HER3 as acceptor (HER3-GFP) was used.

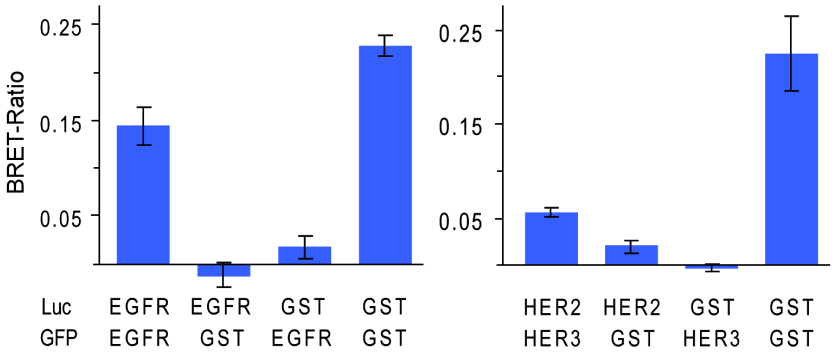

Figure 3.3: Receptor-receptor interaction and GST controls - Receptor-Luc or GST-Luc were transiently coexpressed with either receptor-GFP or GST-GFP in excess in HEK cells. The BRET-ratio of the EGFR pair and the HER2/3 pair were higher than for the negative controls. EGFR: n = 4, HER2/3: n = 2.

First, the BRET signal of both systems was measured. Since part of the receptors is expected to be in the dimeric form under the growing conditions used, detection of a significant BRET-ratio was anticipated. The same control system as in the direct

method (Section 3.2.1, p. 17) with DNA donor:acceptor ratio of 1:20 was used. Indeed, both the EGFR and the HER2/3 pair resulted in BRET-ratios higher than that of the scrambled (receptor-GST) pairs (Fig. 3.3). A competition experiment could not be performed because of the very low expression of the unlabelled receptor (EGFR-Stop, data not shown) but the available controls were convincing enough to move to the analysis of ARNO effect.

3.2.2.1 Influence of ARNO on the receptors BRET-ratio

In this experiment, the effect of coexpression of ARNO in the above systems was investigated. Figure 3.4 shows that ARNO increased the BRET-ratio of both receptor pairs. No effect was detected on the GST pair. The GEF inactive ARNO-E156K mutant did not influence the EGFR BRET-ratio (Fig. 3.5).

Since BRET is very sensitive to both distance and orientation changes, the increase of BRET signal could be due to both enhanced dimerisation or changes in the relative conformation of the receptors. To get a first hint on the mode of action of ARNO, its impact on the constitutive dimeric RTK Insulin-Like Growth Factor Receptor (IGFR) was tested.

Coexpression of ARNO with the IGFR BRET pair lead to an increase of the BRET-ratio (Fig. 3.6), as already observed for the EGFR and HER2/3 pairs. Since the IGFR is constitutively dimerised, this effect can only be attributed to conformational change or receptor clustering. As superresolution light microscopy showed that cytohesin inhibition by SecinH3 does not alter the EGFR cluster size[50], ARNO is most probably not inducing receptor clustering. These results suggest, that ARNO regulates the ErbB receptors by inducing a conformational change of already dimerised receptors.

The analysis above assumes that the expression ratio of RLuc and GFP labelled receptor is constant within a single experiment. In fact, titration of ARNO implies transfection of at least three different plasmids per sample, making it difficult to comply with this condition. An additional difficulty is presented by ARNO activity itself. Indeed, by stimulating ErbB receptor phosphorylation, ARNO contemporaneously induces endocytosis, which can likely change the proportion of donor and acceptor species. SecinH3, as an ARNO inhibitor, offers a possibility to circumvent at least the first problem.

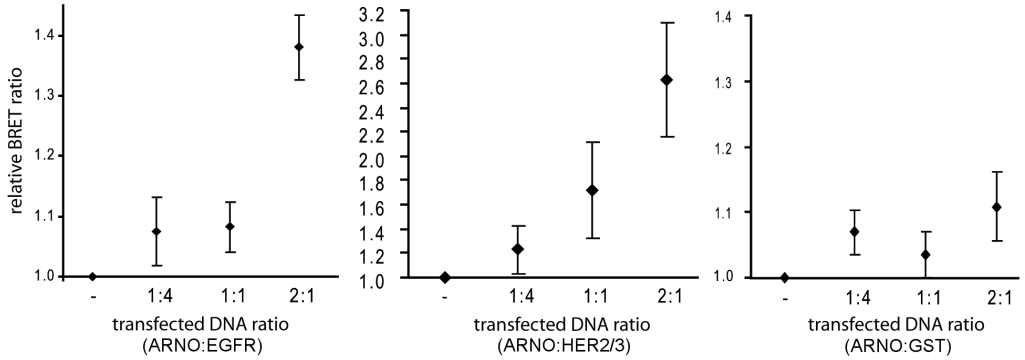

Figure 3.4: ARNO increases the BRET-ratio of the receptor pairs. - Increasing amounts of ARNO DNA were cotransfected with the BRET-pairs DNA in HEK cells. Expression of ARNO enhanced BRET between receptors. No effect was detected on the GST pair. EGFR: n = 11, HER2/3: n = 4, GST: n = 7.

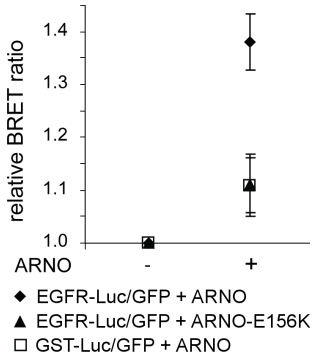

Figure 3.5: ARNO-E156K does not influence the EGFR BRET-ratio - Either ARNO or ARNO-E156K were transfected in 2-times excess over the BRET pair. While ARNO increased by almost 40 % the BRET-ratio of EGFR, the GEF inactive mutant ARNO-E156K had no effect. Indeed the variation of 10 % is comparable to the effect of ARNO on the GST control system. ARNO-E156K: n = 6, ARNO: see Fig. 3.4.

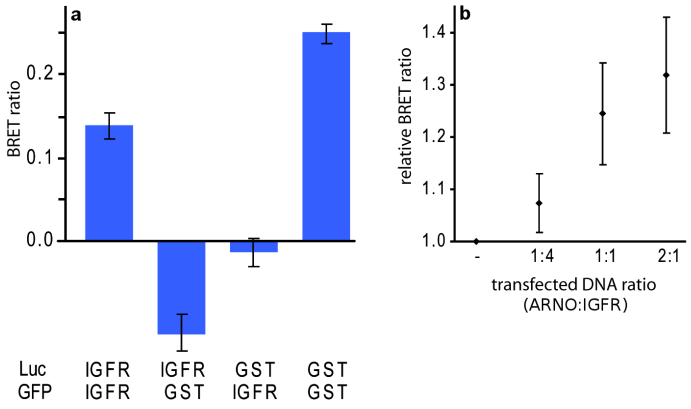

Figure 3.6: ARNO increases the BRET-ratio of IGFR - **a.** IGFR-Luc or GST-Luc were transiently coexpressed with either IGFR-GFP or GST-GFP in excess in HEK cells. The BRET-ratio of the IGFR pair was higher than for the negative controls. n = 1. **b.** Increasing amounts of ARNO DNA were cotransfected with the BRET-pair DNA in HEK cells. Expression of ARNO enhanced BRET between IGFRs. n = 4.

3.2.2.2 Effect of SecinH3

The use of the small molecule SecinH3 allows the use of the same transfection mix for all samples in an experiments, limiting variation in donor and acceptor protein expression ratio. Since SecinH3 was known to be autofluorescent, I had to rule out possible interferences with BRET before using it. From the spectra in Figure 3.7 is clear that SecinH3 is excited and emits at shorter wavelengths than those used in my BRET-system (Fig. 3.1) and the filter set used excludes any interferences. Thus, the small molecule can be used without concern.

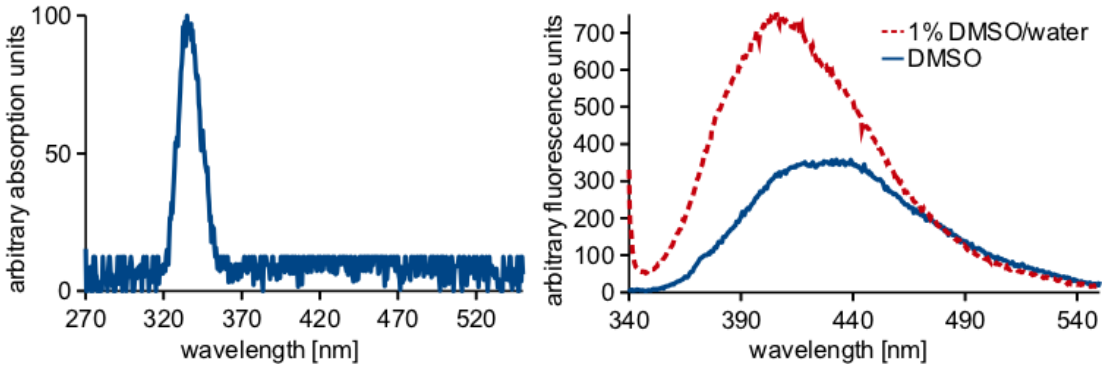

Figure 3.7: SecinH3 absorption and fluorescence spectra - Autofluorescence of SecinH3 does not interfere with BRET measurements (BRET^2: short-wavelength emission at 405 nM, long-wavelength emission at 510 nM). Absorption spectrum was measured with 1 mM SecinH3 in DMSO. Fluorescence spectra were measured with 15 µM SecinH3 in DMSO or 1 % DMSO/water with excitation at 334 nM.

The handling of cells for the BRET measurements was quite different as for our standard western blot analysis, by which inhibition of ARNO by SecinH3 was shown. While for Western blot the cells were stimulated adherent and then harvested on ice, cells for BRET were harvested, transferred into a 96-well plate and stimulated and measured in suspension at 37°C. Thus, I looked for BRET compatible conditions for which inhibition was detectable by western blot were to be defined. Unfortunately, no such condition was found and BRET inhibition with SecinH3 was not achieved.

3.2.2.3 Effect of stimulation on BRET signal

An important experiment to validate my BRET system, was the analysis of the effect of receptor stimulation on the BRET signal. Indeed, the EGF induced dimerisation of EGFR, should be manifest as an increase of BRET-ratio. Thus, HEK cells were

transfected with RLuc and GFP labelled EGFR, harvested after 36 hours and transferred into a microplate. EGF was added to a final concentration of 10 nM 5 minutes before DeepBlueC injection and measurement. Since no effect of the EGF stimulation was detected, various transfection amounts, EGF concentrations and stimulation times were tested, but no condition gave the expected results.

The experiments described above, were all performed with EGFR DNA constructs lacking the exon 4, which codes for an extracellular fragment of the receptor, part of the EGF binding domain[51,52]. A possible explanation of my results was therefore that the receptor constructs used had impaired EGF binding and were therefore not stimulatable. To check this hypothesis, new expression vectors coding for the complete EGFR were cloned (EGFRwE4). Western blot analysis of stimulation showed indeed that the new constructs had a much higher phosphorylation level after stimulation (Fig. 3.8). The stimulation dependent phosphorylation detected up to then in cells transfected with the original constructs, was probably to ascribe to endogenous receptor.

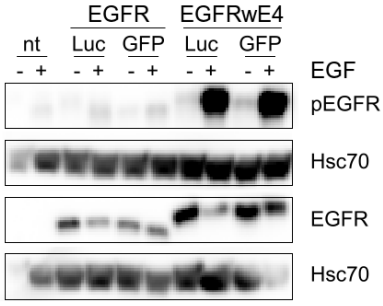

Figure 3.8: Phosphorylation of the EGFR constructs lacking exon 4 is impaired - HEK cells were transfected with expression plasmids of both EGFR isoforms. Stimulation was performed with 10 nM EGF for 5 min. After separation on a 7.5 % SDS-PAGE and Western blotting, the phosphorylated receptor was detected with phosphospecific antibodies. Hsc70 was used as loading control.

The analysis of the effect EGF stimulation was then repeated with the EGFRwE4 BRET pair. The BRET-ratio of aliquots of transfected cells was measured at increasing stimulation times. A reduction of BRET-ratio was detected some seconds after stimulation (Fig. 3.9**a**), but analysis of the cells expressing EGFRw4-Luc only revealed high variability of the background BRET-ratio (Fig. 3.9**b**), undermining the results in **a**. Indeed it turned out, that EGF stimulation affected luciferase activity in the EGFRwE4-Luc receptor. Since this effect was not seen with GST-Luc, it is probably due to a conformational change of the receptor which influences substrate uptake, or variation in the physiological milieu of the luciferase for example because of receptor internalisation. In this set up, as in the direct strategy, I was generally dealing with very low GFP signals. Thus, changes in bioluminescence had a relevant influence on

the GFP signals, altering the resulting BRET-ratios. All this, together with the points discussed in the next section, made interpretation of the data unreliable. Experiments in this direction were therefore temporarily discontinued.

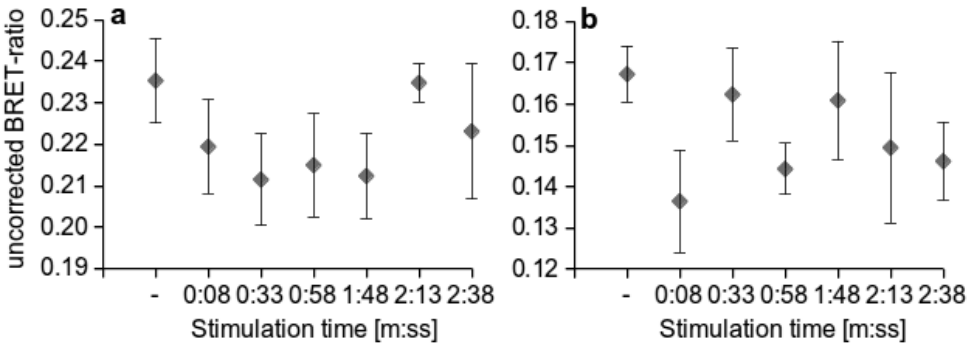

Figure 3.9: Effect of stimulation on BRET-Ratio - HEK cells were transfected with either the EGFRwE4 BRET pair (**a**) or EGFRwE4-Luc alone (**b**). The BRET-ratio was measured at the indicated time after EGF addition and is plotted without subtraction of the background BRET-ratio, to allow visualisation of the results of the donor only samples. A decrease in BRET-ratio in the first minutes after stimulation is visible in **a**, but variability in the same magnitude was detected in cells expressing the donor only (**b**). n = 5.

3.2.3 Discussion

With the intent to elucidate the mechanism by which ARNO activates the EGFR, BRET between EGFR-Luc and ARNO-GFP was analysed. Using GST-Luc and GST-GFP as controls, it was shown that the BRET signal measured in cells coexpressing the EGFR/ARNO BRET pair, was not result of specific binding (Fig. 3.2). Yet, in our recently published work[50], we showed that ARNO directly interacts with the EGFR. The discrepancy between these results is explained by experiments conducted in parallel to the BRET assay, which showed that the ARNO-GFP fusion protein was not functional (A. Bill, J. Theis, unpublished results). To my knowledge the use of ARNO-GFP is not reported in literature. However, various groups successfully worked with the N-terminal chimera GFP-ARNO[29–31,53–59]. Thus, BRET experiments with GFP-ARNO could possibly allow to detect binding of ARNO to the EGFR.

Analysis of the results was complicated by the very low signals detected. As for GFP, it is possible that the problem would be solved by the use of GFP-ARNO, since interaction between the BRET pair is expected to increase the GFP signal. However,

the RLuc signal can only be improved by increasing the transfection amounts or the number of cells per sample. Unfortunately, none of these options was applicable with the experimental conditions of this project.

Competition experiments, in which unlabelled protein is coexpressed with the BRET-pair, are often used as a control of specificity[49]. Yet, for my system the control was unsatisfying for two main reasons. First, I could transfect the cells with only very limited amount of untagged protein DNA and, second, transfection with up to four different plasmids lead to variable ratio of RLuc- and GFP-protein expression. The use of cells stably expressing the donor and acceptor proteins would minimize these problems. Possibly, higher expression of the unlabelled protein could be achieved and, at the same time, the expression ratio of the two BRET-proteins should be less variable. Because the establishment of stable cell lines is particularly work intensive, this would be an interesting option only if further experiments are planned with the cell line.

As a second strategy, an indirect interaction analysis was performed. In this set up, donor and acceptor proteins were both receptor chimeras and the effect of ARNO overexpression was monitored. In both the EGFR-Luc/GFP and HER2-Luc/HER3-GFP expressing cells BRET-ratios higher as for the controls were measured. Interestingly, the EGFR BRET-pair gave a BRET signal almost 3-times higher than that of the HER2/HER3 pair (Fig. 3.4) and similar to that of the constitutive dimeric IGFR (Fig. 3.6). Even if the absolute BRET-signal is usually not comparable between different BRET-pairs (since energy transfer is dependent also on the relative position and orientation of the fluorophores dipoles, which vary from pair to pair) the structural similarity of the receptors studied is tempting to speculation. Indeed, the finding that the EGFR isoform used can actually not be stimulated (Fig. 3.8) because it lacks the exon 4, which is involved in EGF binding, suggests that the energy transfer observed could be a measure of dimerised EGF receptor. This hypotesis is supported by a very recent publication which reports identification of an exon 4-deletion variant of EGFR in gliomas, ovarian cancer tissues and prostate cancer tissues[52]. This mutant displayed minimal EGF binding activity and underwent ligand-independent autophosphorylation and self-dimerisation.

Founding on these results, it is interesting to note that overexpression of ARNO increased BRET between the EGFR pair. This implicate an effect of ARNO on already dimerised receptors, as we could show later with other methods[50]. Nevertheless, these

results must be taken with caution, since they are only valid under the assumption that the expression ratio of RLuc and GFP labelled receptor is constant within a single experiment. In fact, this experiment again requires transfection of at least three different plasmids, a condition which can lead to variations in expression ratio. Moreover, the activity of ARNO itself affect EGFR expression, since activation of the receptor by ARNO overexpression enhances endocytosis.

Experiments were performed to analyse the effect of EGFR stimulation on BRET between EGFR-Luc and EGFR-GFP. With the current knowledge about the characteristics of the exon-4 deletion variant of the EGFR (Fig. 3.8 and Ref. 52), it is clear that no stimulation influence could be detected. With respect to the new constructs EGFRwE4, which were shown to be stimulatable, the analysis is more complicated. Indeed, at first glance, one could interpret the results in Figure 3.9**a** as stimulation dependent change of BRET-ratio. The fact that the ratio is decreasing and not increasing, as one would intuitively expect, can be explained by conformational changes in the EGFR C-terminus, and consequent changes in the fluorophores relative orientation, which overwhelm the effect of dimerisation (a similar result was shown by Yang *et al.* in a luciferase fragment complementation assay[60]). In fact, analysis of the raw data showed a high variability in the luciferase and GFP signals upon EGF stimulation without detectable trends. Again, having to deal with very low signals increased the gravity of these random variations. Additionally, it was observed that the RLuc signal observed for aliquots of the same cell population decreased rapidly with the time. This was caused by deterioration of the DeepBlueeC solution in the instrument. Thus, for comparable luminescence, repeated priming of the instrument was needed during a measuring series. Still, the BRET^2 system used was not adequate for the analysis of stimulation. Ideally, one would monitor the changes of BRET after stimulation in a single sample. Unfortunately, luminescence induced by DeepBlueC is very short lived and allows only a single measurement per sample. The use of BRET^1 or eBRET, in which the substrates Coelenterazine h or EnduRen, respectively, allow detection for up to one to several hours[49], would therefore be more appropriate.

4

Surface plasmon resonance

As an alternative method to investigate a possible interaction of cytohesins and ErbB receptors, surface plasmon resonance (SPR) bioanalysis was chosen because it offers the chance to follow binding events in real time and thus determine a range of interaction characteristics, like association and dissociation rates. These experiments are described in Section 4.3.1.

The small molecule SecinH3 allowed the identification of new roles of cytohesins[17,50], which, in turn, generated interest in SecinH3 derivatives with improved activity and solubility. In Section 4.3.2, the establishment of an SPR based platform for the analysis of the binding properties of these new compounds is described.

Although SPR has been used since the early Nineties for the analysis of biomolecular interactions, still various misconceptions are diffused. Thus, I will start with a detailed introduction to the theory of SPR (Section 4.1) and its application in affinity biosensors (Section 4.2).

4.1 Physics of surface plasmon resonance

Surface plasmon resonance is a phenomenon that occurs in planar metal-dielectric waveguides, as for example a metal/water interface[61]. Surface plasmons are electrons oscillations which propagate parallel to the interface and are usually generated by means of a prism coupler and the attenuated total reflection method[61]. Our device is using the Kretschmann geometry, in which the metal film is evaporated directly onto the prism[61,62]. When illuminated, the metal film reflects part of the light back into the

prism while a part of the light propagates in the metal as a so-called evanescent wave (an inhomogeneous electromagnetic wave which decays exponentially in the direction perpendicular to the prism-metal interface)[61]. If the metal film is sufficiently thin, the evanescent wave penetrates through it. For each wavelength, a single angle of incidence leads to excitation of surface plasmons at the outer boundary of the metal film via the evanescent wave field. As a result, SPR is seen as a drop in the intensity of the reflected light (Fig. 4.1)[61,62].

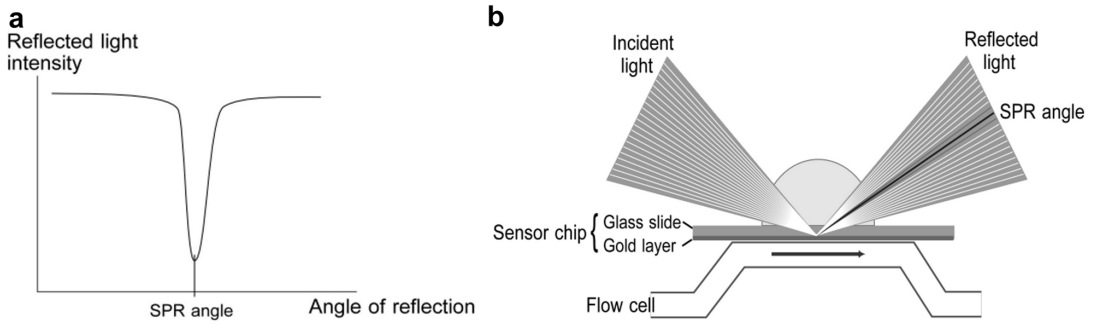

Figure 4.1: The SPR angle - At a certain combination of wavelength and angle, the incident light excites plasmons in the gold film. As a result, a characteristic absorption of energy via the evanescent wave field occurs and SPR is seen as a drop in the intensity of the reflected light.[62] *Reprinted from* Biacore - Sensor surface handbook *(Ref. 62), copyright 2005-2007 GE Healthcare Bio-Sciences AB.*

Application of the perturbation theory to the electromagnetic theory of optical waveguides demonstrate that the propagation of surface plasmons is highly sensitive to changes in the refractive index at the boundary[61], such as those resulting from adsorption of molecules to the metal surface. This is the property which is exploited in optical sensors based on surface plasmons[63], such as Biacore[62].

Our instrument, a Biacore 3000, is an SPR sensor with angular modulation. That is, it uses monochromatic light for excitation and monitors the reflected light intensity at multiple angles of incidence. Resonance is thus detected as a dip in the intensity of the reflected light (Fig. 4.1**a**) and the sensor output is the angle of incidence yielding the lowest intensity (Fig. 4.1**b**)[62,63]. Changes in this angle are expressed in Resonance Units (RU).

4.2 SPR affinity biosensors

Since the Nineties commercial SPR based sensor devices for the analysis of biomolecular interactions are available. SPR affinity biosensors carry biorecognition elements which are able to interact with a selected analyte. Binding of analyte molecules to the affinity elements produce an increase in the refractive index at the sensor surface an thus a change in SPR[63].

The first and best known commercial SPR instrument is the Biacore, commercialised by Pharmacia Biosensor AB (now GE Healthcare Bio-Sciences AB). For this project a Biacore 3000 was used, whose principal characteristics are described in the next section.

4.2.1 The Biacore system

The principal parts of Biacore 3000 are the detector unit, an interchangeable sensor chip and the microfluidic system. The sensor chip is composed of a glass slide with a thin layer of gold where SPR take place. The gold surface is derivatised, as described in Section 4.2.3.1 (p. 31), to allow immobilisation of the biorecognition element. The sensor chip is pressed against a semi-cylindrical glass prism and light from a diode is focused on to the sensor surface, covering a fixed range of incident and reflected angles[62]. On the opposite side, the chip is pressed on the flow cell block, so as to form a wall of the flow cells and the gold layer is thus in direct contact with the flowing buffer (Fig. 4.1**b**). The four flow cells allow analysis of up to three ligand in parallel (the fourth cell is usually used as reference to correct for bulk refractive index changes in the sample). The flow cell block is part of the Integrated Microfluidic Cartridge (IFC), which consists of a series of channels and valves and, together with the pump system, is responsible for precisely controlling the delivery of sample and buffer. The use of two pumps assure a continuous buffer flow even during sample preparation and injection[62,64].

4.2.2 Measuring approach and terminology

To enhance comprehension, some basic terminology is explained here. In Biacore experiments, one interacting partner (the **ligand**) is immobilised on the surface of a sensor chip while a solution with the second interaction partner (the **analyte**) is flown over

the surface along four **flow cells** (Fc). The sample is carried in a continuous flow of buffer, the **running buffer**.

The **resonance units** (RUs) quantify the response, which is directly proportional to the concentration of biomolecules on the surface and is acquired in real time. The progress of the interaction is displayed on a **sensorgram**, a plot of response against time. In a typical **cycle** a **baseline** is collected while running buffer is flowing. Then, **association** is observed during sample injection, followed by **dissociation** when running buffer is again flown on the chip. **Regeneration**, that is removal of bound analyte from the surface without damaging the ligand, is achieved by injection of a regeneration solution with composition dependent from the investigated interaction. Typical regeneration solutions are high salt and acidic buffers and detergent containing solutions. After regeneration a new cycle with a different analyte concentration is measured. The sensorgram of a typical cycle is displayed in Figure 4.2, but before discussion of the most important characteristics of a simple binding curve, some details of the experimental design are presented.

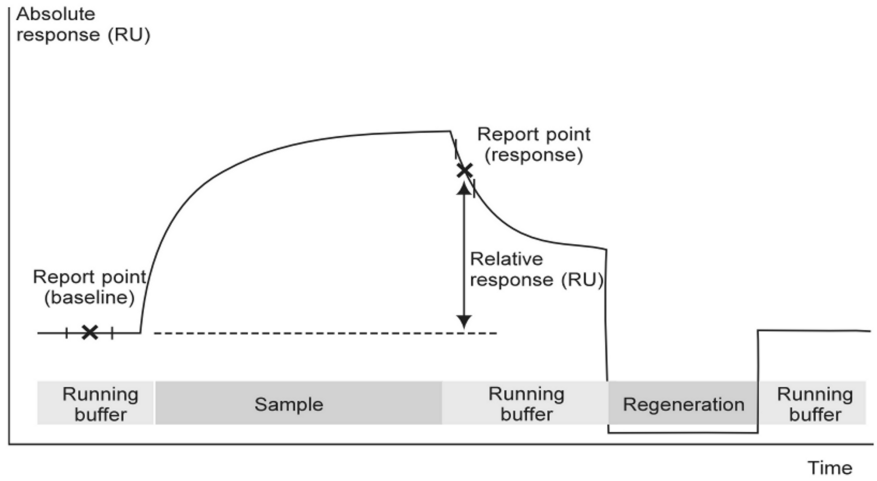

Figure 4.2: Schematic illustration of a sensorgram - The bars below the sensorgram curve indicate the solutions that pass over the sensor surface. *Reprinted from* Biacore - Sensor surface handbook *(Ref. 62), copyright 2005-2007 GE Healthcare Bio-Sciences AB.*

4.2.3 Experimental design

Acquiring SPR data is easy, *generating high quality data is an art form and takes some effort*[65]. In particular, it is important to start with good quality reagents (this include

chemically and conformationally pure ligand and analyte, since *in order for data to fit a simple bimolecular reaction model, the analyte and ligand must be monomeric in solution and form a 1:1 complex when mixed*[66]). A critical point is selecting the interaction partner to immobilise and the immobilisation procedure, as discussed below.

4.2.3.1 Immobilisation

Although initially immobilisation was based on simple physical adsorption of proteins to an active metal surface, unmodified gold is usually not a suitable surface for biomolecular interactions. This is because of its high tendency for adsorption of proteins and other molecules, the possible loss of activity of these molecules upon adsorption, and the instability of this type of surface binding[67]. Thus, to protect the biological sample from contact with the gold and at the same time provide a mean of attachment of a surface matrix, the gold on almost all Biacore sensor chips is covered with a self-assembled monolayer of alkanethiol molecules[62,67].

On most sensor chips, such as the CM5/CMDP*, the surface is additionally covered with a matrix of carboxymethylated dextran, a flexible unbranched carbohydrate polymer forming an approximately 100 nm thin hydrogel-like polymer layer which provides high flexibility and water solubility. Additionally, the dextran matrix offers a defined chemical basis for covalent attachment of biomolecules to the surface and the negatively charged carboxyl groups allow electrostatic concentration of positively charged molecules, enabling efficient immobilization from dilute ligand solutions[62] (see Section 4.5, p. 33). Moreover, the three-dimensional structure increases the binding capacity several-fold in comparison with a flat surface and extends the region where interactions occur to match with the penetration depth of the evanescent wave[62,67].

Carboxymethylated dextran derivatised chips are compatible with all the three main strategies for immobilising biomolecules on the sensor chip surface shown in Figure 4.3, namely covalent coupling, non-covalent capture and lipid layers mediated coupling.

With respect to covalent coupling, the carboxylic acid residues of the carboxymethylated dextran derivatised chips can be either reacted with amines or other nucleophilic groups, or converted for use in coupling chemistries based on, e.g., aldehyde and carboxylic acid condensations and thiol reactions. In fact, the formation of a covalent

*Sensor chips were bought from two companies. The first name always refers to the Biacore chips and the second one to the Xantec chips (see *Materials and methods*, p. **??**).

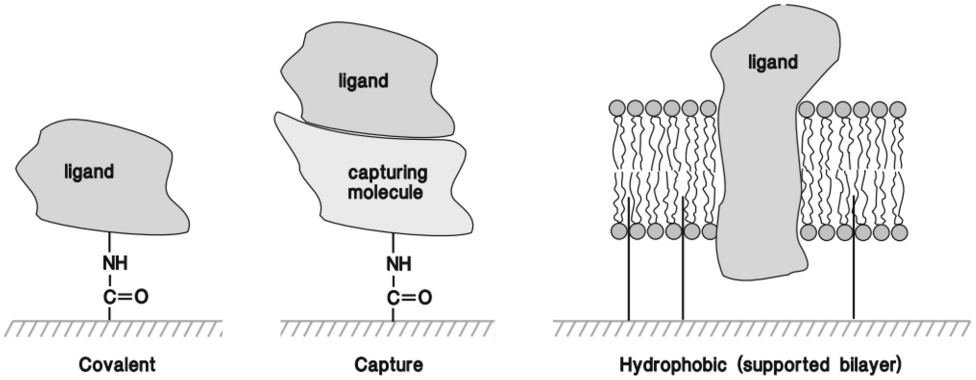

Figure 4.3: Strategies of biomolecules immobilisation - The ligand can either be chemically coupled to accordingly derivatised surfaces or immobilised by binding to a capture molecule such an antibody or streptavidine. As a third method, hydrophobic carrier such as a lipid monolayer or bilayer can be attached to the sensor chip surface through hydrophobic adsorption and offer, e.g, a matrix for insertion of membrane-associated proteins. *Reprinted from* Biacore - Sensor surface handbook *(Ref. 62), copyright 2005-2007 GE Healthcare Bio-Sciences AB.*

amide bond between the carboxylic acid and the amine group in lysine residues is the most common immobilisation method in Biacore experiments[67]. The coupling is usually performed directly in the instrument and in two steps: the carboxylic group is activated with 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) before the ligand is passed on the surface, to avoid reaction between the carbodiimide and the ligand. Since the reactivity of the first step intermediate in aqueous solutions is so high that water hydrolysis would rapidly transforms it back to carboxylic acid, a mixture of EDC and *N*-hydroxysuccinimide (NHS) is used to form a more stable active ester derivative (Fig. 4.4)[67].

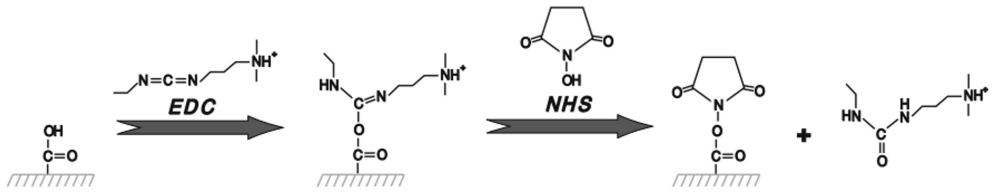

Figure 4.4: Activation of carboxymethyl dextran with EDC/NHS - The carboxylic group is activated with EDC. To avoid rapid hydrolysis, NHS is present in the coupling solution to give reactive succinimide esters which react spontaneously with amines and other nucleophilic groups. For simplicity the dextran chains are omitted. *Reprinted from* Biacore - Sensor surface handbook *(Ref. 62), copyright 2005-2007 GE Healthcare Bio-Sciences AB.*

Traditionally, coupling of active esters is done under slightly alkaline conditions where a normal alkylamine nucleophile is close to its pKa and can compete with water hydrolysis and this is the preferred method for coupling of organic molecules and small peptides[67]. However, for immobilization of proteins, where high-density modifications are desirable, electrostatic preconcentration is usually used. This method takes advantage of the unreacted carboxylic groups on the EDC/NHS activated surface. As these leads to an overall negatively charged environment, under low ionic strength light acidic buffer conditions positively charged proteins are attracted resulting in a high local surface concentration, which in turn favours protein coupling over water hydrolysis (Fig. 4.5)[67]. This methods leads to high-density surfaces and has the additional advantage that coupling occurs under conditions where only a small fraction of the nucleophilic groups on the protein are reactive and thus the very few immobilization points increase the likelihood of preserving activity[67]. Additional methods of covalent coupling include the formation of disulfide bonds and thioether linkages, and coupling to aldehyde groups and will not be discussed here.

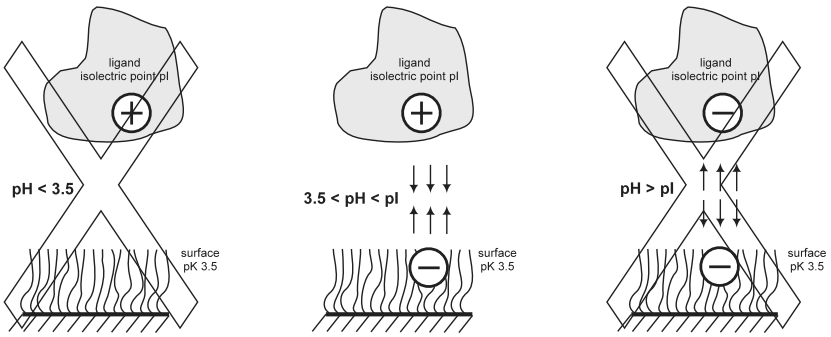

Figure 4.5: Preconcentration of the ligand - The ligand is concentrated on the surface through electrostatic attraction when the pH lies between the isoelectric point (pI) of the ligand and the pKa of the surface. If the pH is too low or too high, the ligand will not be concentrated on the surface[62]. *Reprinted from* Biacore - Sensor surface handbook *(Ref. 62), copyright 2005-2007 GE Healthcare Bio-Sciences AB.*

Where covalent coupling does not work, capture-based coupling can be used. In this technique a capture molecule, often an antibody, with high affinity for the ligand of interest is coupled to the chip and the ligand is then immobilised through binding to the capture molecule. The advantage of this setup is the possibility of removing the ligand after analysis. This step is analyte independent but new ligand must be captured for the next assay cycle, thus higher amounts of ligand are needed. Moreover,

one should take care that the affinity of ligand and capture molecule is sufficient to avoid loosing ligand during the measurement[67].

Antibodies specific for tagged recombinant proteins are often immobilised as capture molecule, but also organic molecule can be used for this scope. An example is nitrilotriacetic acid (NTA) which in complex Ni^{2+} ions is routinely used in affinity purification of polyhistidine tagged proteins[68]. Accordingly, His-tagged proteins can be immobilised by chelation on NTA derivatised sensor chips. A particular advantage of this strategy is the easy removal of ligand by sequestering the nickel ions with chelating agents such as EDTA (ethylenediaminetetraacetic acid).

The third strategy, hydrophobic attachment of membrane-associated proteins through lipid monolayers and bilayers, is not relevant to this project and is therefore not presented in more detail.

4.2.3.2 Data processing

Especially when measuring small signals, a careful data processing is necessary to get rid of system artifacts. In fact, the noise from the injection needle, bulk refractive index changes, washing steps, and instrument drift can be similar in magnitude to the binding signal. If not stated otherwise, we applied *double referencing* (i.e. both the signal of a reference cell, as well as a the signal of blank injections were subtracted from the binding signal) as described by Myzska[66] (see Section 7.1.6.2, p. 133, for more details). Global analysis was then applied, as described in next section.

4.2.4 Data analysis

The ability of SPR biosensors to detect complex formation in real time makes it possible to obtain quantitative information about binding interactions, including the rates of association and dissociation. The binding constants are usually determined by either kinetic or equilibrium analysis. Below, the evaluation models and the requirements to the data are presented.

4.2.4.1 Kinetic analysis

Biacore is best suited for kinetic analysis of simple bimolecular interactions of the type

$$\mathrm{A} + \mathrm{L} \underset{k_\mathrm{d}}{\overset{k_\mathrm{a}}{\rightleftharpoons}} \mathrm{AL}\,, \tag{4.1}$$

where A is the analyte, L the ligand and k_a and k_d the association and dissociation rate constants, respectively. When the number of analyte molecules is higher than the amount of ligands, the surface concentration of the complexes γ formed per unit time t is described by the Langmuir equation for interactions at a surface in contact with reactants in solution (pseudo first-order reaction)[69]

$$\frac{\mathrm{d}\gamma}{\mathrm{d}t} = k_\mathrm{a}\,\alpha_0\,(\beta - \gamma) - k_\mathrm{d}\,\gamma\,, \tag{4.2}$$

where β is the surface concentration of ligands and α_0 the analyte initial concentration. Its solution is a single exponential function with an asymptote corresponding to the equilibrium fulfilling equation[69]

$$K_\mathrm{A} = \frac{k_\mathrm{a}}{k_\mathrm{d}} = \frac{\gamma_\mathrm{eq}}{\alpha_0\,(\beta - \gamma_\mathrm{eq})}\,. \tag{4.3}$$

This condition is respected by Biacore 3000, where the concentration of free analyte is controlled by the solution flown in the cell and can be either increased or decreased stepwise. Thus, an SPR sensorgram encompasses an association phase (beginning with the sharp increase of the free analyte concentration to a constant value α_0) followed by a dissociation phase (when the free analyte concentration is suddenly reduced to zero)[69]. Both association and dissociation phase have an exponential progress, as shown in Figure 4.6**A-F**.

The best way to accurately estimate the binding constants is to perform global analysis. That is, all responses of a data set (i.e. responses of a wide range of concentrations) are fitted simultaneously using the same set of constants (k_a, k_d, γ_max). The fitting procedures of analysis softwares like BiaEvaluation, base on minimization of the chi-squared value χ^2, given by the difference between simulated and experimental data. For this, the initial parameter estimates are adjusted iteratively as long as χ^2 can be improved. As global analysis is particularly stringent, it provides a method to discriminate between different reaction models[70].

The model described above premises uniform analyte concentration distribution in the flow cell. In fact, the concentration profile is dependent on analyte diffusion and binding. When the binding rate is equal or faster than the diffusion of analyte to the ligand surface, mass transport must be considered in the evaluation model[70].

Mass transport limitation

A 1:1 binding reaction influenced by mass transport can be described by the equation

$$\mathrm{A}_0 \underset{k_\mathrm{M}}{\overset{k_\mathrm{M}}{\rightleftharpoons}} \mathrm{A} + \mathrm{L} \underset{k_\mathrm{d}}{\overset{k_\mathrm{a}}{\rightleftharpoons}} \mathrm{AL} \tag{4.4}$$

with the mass transport coefficient k_M characterising the rate of analyte A_0 diffusion to and from the reaction surface[70].

The effect of mass transport limitation on a typical sensorgram is shown in Figure 4.6**F**. Since slow analyte transport causes a concentration decrease when it is bound during the association phase and an increase when it is released during the dissociation phase, both reactions are slowed down[69]. Typically, mass transport introduces a linear component at the beginning of the injection phase and makes the dissociation phase non-exponential[65].

Since laminar flow, as found in a Biacore flowcells, is very ineffective in transporting analyte to the sensor surface, translational diffusion (i.e diffusion along a concentration gradient, in this case perpendicular to the chip surface) is particularly important in the vicinity of the active sensor layer[69]. The diffusion coefficient D, which expresses the proportionality between the rate of diffusion and the concentration gradient, decreases as the size of the molecule increases and is dependent on temperature and viscosity of the medium. For typical biomolecules in aqueous medium, D is usually between 10^{-7} cm^2 s^{-1} and 10^{-6} cm^2 s^{-1} (Ref. 69).

In the simplified model of mass transport limitation for a pseudo first-order reaction, the average concentration of complexes $\langle\gamma\rangle(t)$ is described by

$$\frac{\mathrm{d}\langle\gamma\rangle(t)}{\mathrm{d}t} = k_\mathrm{a}^\mathrm{ef}(t)\alpha_0[\beta - \langle\gamma\rangle(t)] - k_\mathrm{a}^\mathrm{ef}(t)\langle\gamma\rangle(t)\,, \tag{4.5}$$

where the effective rate constants

$$k_\mathrm{a}^\mathrm{ef}(t) = \frac{k_\mathrm{a}}{1 + k_\mathrm{a}[\beta - \langle\gamma\rangle(t)]/k_\mathrm{M}} \quad \text{and} \quad k_\mathrm{d}^\mathrm{ef}(t) = \frac{k_\mathrm{d}}{1 + k_\mathrm{a}[\beta - \langle\gamma\rangle(t)]/k_\mathrm{M}} \tag{4.6}$$

are both space- and time-dependent and

$$k_{\mathrm{M}} \approx 1.378 \left(\frac{v_{\max} D^2}{hl} \right)^{1/3} . \tag{4.7}$$

Again, α_0 and β are the analyte initial concentration and the ligand surface concentration, respectively. The maximum flow velocity v_{max} is found at the mid-point of the cell height h (dimension perpendicular to the sensor surface) while l is the length of the cell (dimension along the flow)[69].

The effect of k_{M} on the SPR response is estimated by the Damköhler number

$$Da = k_{\mathrm{a}} \beta \left(\frac{v_{\max} D^2}{hl} \right)^{-1/3} \sim \frac{k_{\mathrm{a}} \beta}{k_{\mathrm{M}}} , \tag{4.8}$$

which is proportional to the ratio of the reaction velocity to the diffusion flux of the analyte at the beginning of the association phase[69].

For $Da \ll 1$ the mass transport is much faster than the surface reaction itself and its effect can therefore be ignored, while for $Da \gg 1$ the reaction is completely controlled by diffusion and is thus not possible to determine reaction rate constants from the sensorgram[69].

From Equation 4.8 is evident that both increasing $v_{\max}$ and decreasing the amount β of coupled ligand reduce the Damköhler number. This is the motivation of the general advice to use low ligand density surfaces and high flow rates when determination of kinetics constant is the goal of an SPR experiment.

4.2.4.2 Equilibrium analysis

If the association phase is long enough, steady response levels are reached, visible as plateau at the end of injection. If this is true for all concentrations (as it is the case in Figure 4.6**C** and **D**), the whole data set can be used in an equilibrium analysis to determine affinity[65].

From the reaction in Equation 4.1 and the equilibrium dissociation constant definition

$$K_{\mathrm{D}} = \frac{k_{\mathrm{d}}}{k_{\mathrm{a}}} = \frac{\alpha_{\mathrm{eq}}\, \beta_{\mathrm{eq}}}{\gamma_{\mathrm{eq}}} \quad \text{and with} \quad \gamma_{\max} = \beta_{\mathrm{eq}} + \gamma_{\mathrm{eq}} \tag{4.9}$$

the relation between steady state binding level and analyte concentration

$$\gamma_{\mathrm{eq}} = \frac{\gamma_{max}}{(1 + K_{\mathrm{D}}/\alpha_{\mathrm{eq}})} \tag{4.10}$$

is derived. Thus, the K_{D} can be determined by measuring the dependence of the equilibrium response on the injected analyte concentration (binding isotherm)[69].

4.2.4.3 How to recognise a good spectrum

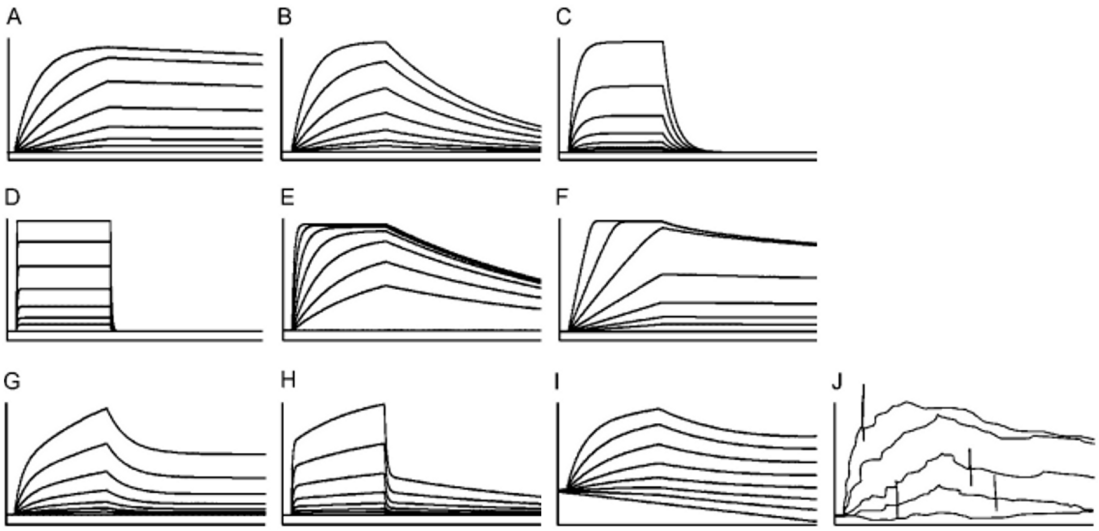

Figure 4.6: Simulated binding responses - Panels A-F depict systems described by simple exponentials and panels G-H depict problematic systems. **A-C.** Interactions having the same k_{a} but increasing k_{d} across the series. **D.** An interaction that reaches equilibrium within a few seconds of the start of the injection phase. **E.** Responses that, at higher analyte concentrations, approach saturation of the ligand surface. **F.** A partially mass transport-limited interaction. **G, H.** Interactions displaying complexity. **I.** Drifting responses. **J.** Severely substandard responses, which often result from analyte aggregation/precipitation and/or instrument maintenance issues[65]. *Reprinted from Journal of Molecular Recognition (Ref. 65, p. 358), copyright 2008, with permission from John Wiley and Sons, Inc.*

Now that the theory of SPR binding analysis is explained, the main features of the sensorgrams in Figure 4.6 should be readily detectable. The sensorgrams **A-C** show typical binding curves where both association and dissociation phase can be described by a simple exponential. Across these three data sets the k_{a} is kept the same but the k_{d} increased 100 fold to demonstrate how the k_{d} affects the binding profile in both the association and dissociation phases[65]. The square-shaped response in **D** is result of

a very fast k_d but can still be described by the standard mechanism in Equation 4.1. The same is true for the data set **E**, where at high concentrations the analyte start saturating the ligand surface[65].

It is important to distinguish saturation from equilibrium. Indeed, saturation is observed when the ligands on the surface are fully occupied with analyte. Under conditions that approach saturation, the responses for increasing analyte concentrations have nearly the same intensity at the end of association. This maximal response R_{max} can be used to determine the binding capacity of the surface and estimate the activity of the immobilized ligand[62,65]. However, often the response continue to increase together with analyte concentration. This non-specific binding is mostly due to heterogeneity of the interaction partners and should not be interpreted as second type of binding. Measurements should preferably be performed around the K_D ($\sim$ 0.1 to 10 fold) as there is no absolute necessity to demonstrate saturation in a biosensor experiment[65].

The sensorgrams in **F** are an example of a partially mass transport limited reaction. Particularly evident is the linear initial binding rate and the non exponential dissociation phase. As long as the binding reaction is not completely covered by the mass transport effect, the kinetic parameters can easily be determined by application of the right model (see Section 4.2.4.1). Nevertheless it is advisable to repeat the measurements with lower surface density and higher flow rate to avoid mass transport effects[65].

Data sets **G** and **H** show complex binding profiles. This curves can not be described by a simple exponential and are commonly referred as biphasic binding response. Although this behaviour could be biologically relevant[71], *the fact is that once you get biphasic data it becomes virtually impossible to resolve what actual event is leading to the complex response. This is because biphasic data can be described equally well by a multitude of models.*[65] Of course a more complicate model with an higher number of variable will fit the data better, but this does not mean that the model is appropriate. Therefore before trying to find a model which perfectly fit the results obtained, it would be better to redesign the assay to see if one can get rid of the complex part[65].

As last, **I** and **J** are examples of inappropriate experimental conditions. The drifting baseline in **I** indicate inefficient coupling, as the ligand is detaching from the sensor surface during the measurement, while the sensorgrams in **J** are a typical example of measurement with a dirty instrument or bad quality reagents[65].

4.3 Results and discussion

As described above, to determine reliable kinetics parameters out of an SPR sensorgrams set, high quality data, ascribable to a simple binding mechanism, are required. Often, a careful evaluation of various coupling and buffer systems is required to get data of the desired quality. The next sections give an overview of the optimisation process for the analysis of ARNO/EGFR (Section 4.3.1) and ARNO/small molecule (Section 4.3.2) interaction, describe some common challenges and present the obtained binding results.

4.3.1 Analysis of ARNO/EGFR binding affinity

To start with, interaction of ARNO and EGFR was tested. Since expression and purification of the receptor is quite more demanding, we preferred it as immobilised partner, to limit the amount needed for analysis. Immobilisation on NTA/NiP chips was the first method tested, because it allows loading of new, native, ligand whenever necessary.

4.3.1.1 Immobilisation on nitrilotriacetic acid derivatised chips

A typical analysis cycle on NTA/NiP chips consists in loading the NTA with Ni^{2+} by flowing a $NiCl_2$ solution on the flow cells (Fc). Then, the His-tagged protein binds to the surface by chelation of the Ni ions. Binding analysis by injection of analyte is followed by regeneration with either EDTA (which remove the ligand together with the analyte, due to depletion of Ni^{2+}) or an appropriate regeneration solution (which remove the bound analyte without affecting the ligand activity).

To get acquainted to the system, a protein pair known to interact was chosen. Thus, interaction between His-ARNO-Sec7 (the purified Sec7 domain of ARNO carrying an His-tag) and ARF6 carrying a strep-tag[72] (strep-ARF6) was measured.

Interaction of ARNO and ARF6

First, the loading conditions were tested. All flow cells were rinsed with EDTA to get rid of Ni^{2+} and other contaminant ions and single Fcs were loaded with $NiCl_2$ and at the end the protein solution flown over all cells. Fc without Ni^{2+} were used to check

the requirement of Ni^{2+} for His-tag protein immobilisation. In fact, some His-ARNO-Sec7 was found to bind also in the Ni-free Fc 1 (Fig. 4.7). The initial assumption of a damage of the IFC, leading to diffusion of Ni^{2+} from Fc 1 to Fc 2, was discredited by a system check which did not reveal any particular IFC failure. The effect was thus ascribed to carry-over of $NiCl_2$ after injection.

Two main tricks were applied to reduce the problem: 3 mM EDTA was added to the dispenser buffer (connected to pump B and used for sample preparation and washing of the injection needle) and another rinsing program was used. Indeed, the command 'WASH IFC' leads to increased risk of carry-over. The commands 'EXTRACLEAN' or 'WASH Needle' should be used instead. With this modifications carryover was successfully avoided.

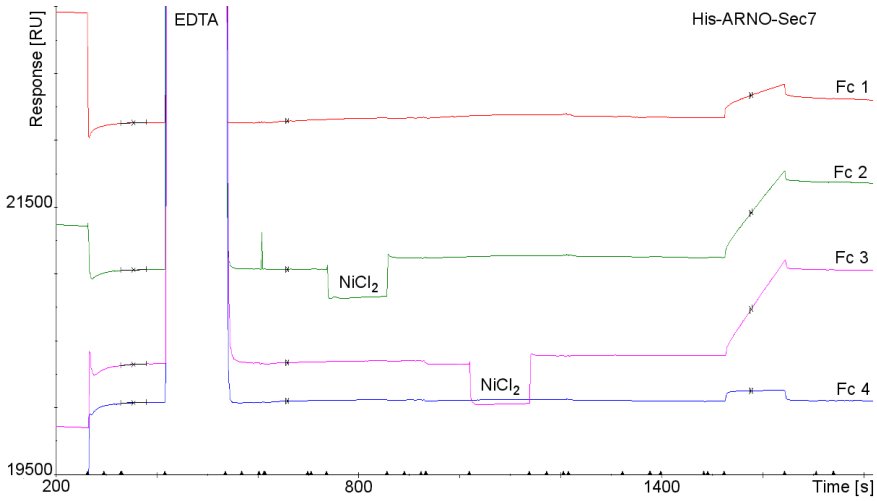

Figure 4.7: Carry-over of Ni ions leads unspecific binding of His-ARNO-Sec7 to a Ni^{2+}-free flow cell - All flow cells were rinsed with EDTA and Ni^{2+} was then loaded on Fc 2 (green) and 3 (magenta) only. His-ARNO-Sec7 bound to Fc 1 (red), too. Fc 4 (blue) was Ni^{2+}-free and no binding was detected. The problem was solved by use of a robuster buffer system and special cleaning procedures. Responses are presented unreferenced.

For binding analysis the cells were loaded with either nothing (Fc 1), Ni^{2+} (Fc 2), Ni^{2+} and His-ARNO-Sec7 (Fc 3), or Ni^{2+} and His-PAZ* (Fc 4). When strep-ARF6 was injected, strong binding to the Ni^{2+}-loaded cell was detected (Fig. 4.8). Metal-dependent, nonspecific protein adsorption to the biosensor surface was already reported

*The purified PAZ domain of Argonaute 1 was used as negative control because it has similar size and like the Sec7 domain carries a His-tag. Expression and purification of His-PAZ are described in Refs. 73 and 74.

by Willard and Siderovski[75]. Analysis is thus only possible when the reference cell has the same surface concentration of protein as the Fc of interest (and comparable levels of unspecific binding to exposed Ni^{2+} is expected). Therefore, I always tried to couple similar amounts (in RU) of proteins of comparable size on the two cells. Subtraction of the signal measured in Fc 4 (immobilised His-PAZ) from that in Fc 3 (Sec7 domain) gave a reasonable binding curve for 2 µM strep-ARF6 (Fig. 4.8**b**).

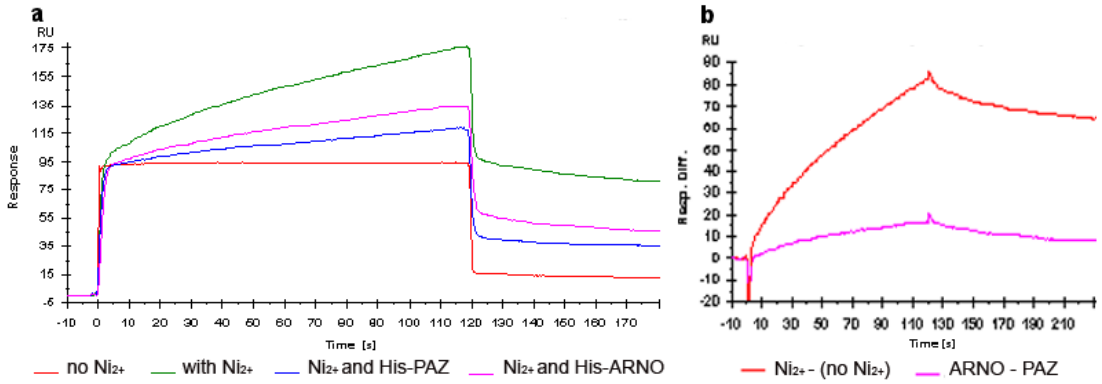

Figure 4.8: Metal dependent unspecific adsorption of non His-tagged proteins to the Ni-NTA surface - This complication was already reported[75]. **a.** The raw response data of four different cells (Ni-free (red) and loaded with Ni^{2+} only (green), His-PAZ (blue) and His-ARNO-Sec7 (magenta), respectively) are shown. The Ni-loaded Fc showed the highest response over injection of 2 µM strep-ARF6. **b.** Referenced curves from **a**: in red the response of the Ni^{2+}-loaded cell after subtraction of the Ni^{2+}-free cell is given. In magenta the response of the ARNO loaded Fc after subtraction of the reference cell with PAZ.

Acquisition of a complete data set showed concentration dependent binding of strepARF6 on His-ARNO-Sec7 (Fig. 4.9). The control proteins GST, Rac1-strep and Rac1-GST* did not bind even at high concentrations. However, the curve shapes did not fit a simple 1:1 binding model. Because the responses also did not reach equilibrium, it was not possible to determine the binding characteristics.

Since analysis of the ARNO-ARF6 interaction was only meant as training, we decided not to invest time in optimisation and to directly test if the EGFR-ARNO interaction was better behaving.

*Rac1-strep and Rac1-GST were kindly provided by Björn Niebel.

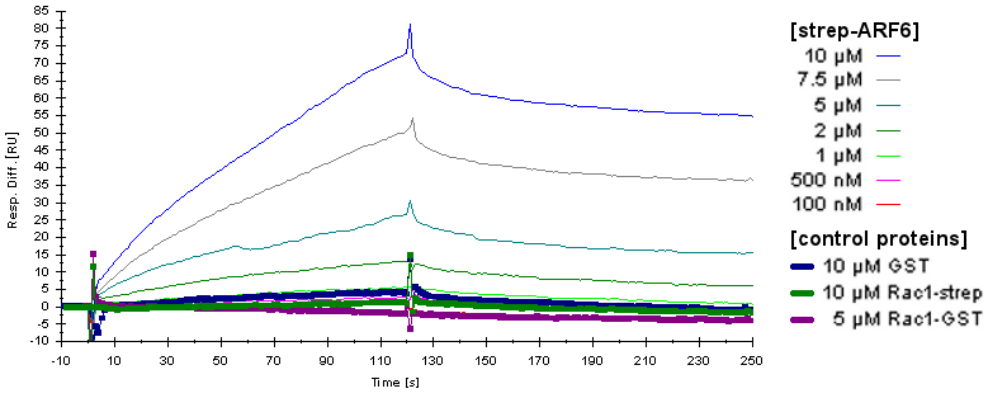

Figure 4.9: StrepARF6 bound immobilised His-ARNO-Sec7 in concentration dependent manner - However, the responses could not be described by a simple binding model. Controls proteins did not show binding. Data were referenced with the His-PAZ signal.

Interaction of EGFR and ARNO

As described above, we were interested in immobilising the EGFR, thus a construct consisting in the intracellular part of the EGFR (EGFR-ICD) with an His-tag (His-EGFR) and a strep-tagged ARNO (strep-ARNO) were used.

His-EGFR was immobilised according to the method developed in the previous section and His-PAZ was again immobilised on the reference cell. The responses obtained by flowing strep-ARNO on the Fc were primarily concentration dependent but, at times, outliers were observed (Fig. 4.10). GST did not bind to the receptor.

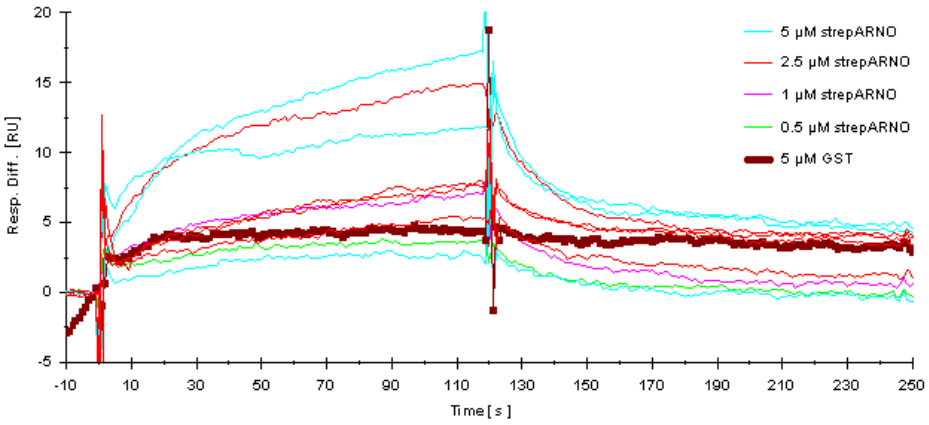

Figure 4.10: Sensorgrams for binding of ARNO to the coupled EGFR - The responses were concentration dependent but outlier were observed. The negative control GST was not binding. Data were referenced with His-PAZ.

The receptor is bigger (> 60 kDa) than the PAZ (19 kDa) domain, which is therefore probably not an optimal negative control in this case (in particular with regards of the problem of unspecific adsorption to Ni^{2+}). Looking for a more appropriate control protein, His-ERK (a 42 kDa His-tagged protein) was found*. Unfortunately, the immobilised His-ERK had very strange behaviour on the chip and could not be used as reference and no better control was available.

Since the unspecific binding issue was difficult to keep under control, I tried direct covalent coupling of the receptor on a CM5/CMDP chip.

4.3.1.2 Preconcentration on carboxymethylated dextran chips

Covalent coupling offers the advantage that, since coupling of the receptor is performed only once, only a minimal amount of His-EGFR is needed and, more important, it avoids the presence of Ni^{2+} and the resulting unspecific binding. However, it as the disadvantage that the ligand must withstand the coupling conditions (e.g. acidic pH) and has to be regenerated after each analyte injection (i.e. one has to find conditions to elute the analyte without denaturing the receptor).

As explained in Section 4.2.3.1, p. 33, proteins are usually coupled at pH comprised between 3.5 and their pI, the range in which electrostatic attraction favours preconcentration on the surface. His-EGFR was thus desalted in 10 mM NaOAc at pH 5.0 and injected on a non activated CMDP chip to test preconcentration. Figure 4.11 shows, that the response did not go back to the basis level in the washing phase: the receptor could only be removed under harsh conditions. Sticking on the chip surface is probably due to denaturation of the receptor under acidic conditions.

Beacuse of this problem, preconcentration was not possible. I thus decided to combine NTA immobilisation with covalent coupling as described in next section.

4.3.1.3 Covalent coupling to nitrilotriacetic acid derivatised chips

NTA/NiP chips are obtained by derivatisation of CM5/CMDP chips. Since EDC/NHS activation is never complete, free carboxyl groups are still present on the surface and can be used to covalently bind the protein via standard amine coupling chemistry. Thus, it is possible to take advantage of the binding of the His-tag to Ni-NTA to concentrate

*His-ERK was kindly provided by Nicole Krämer.

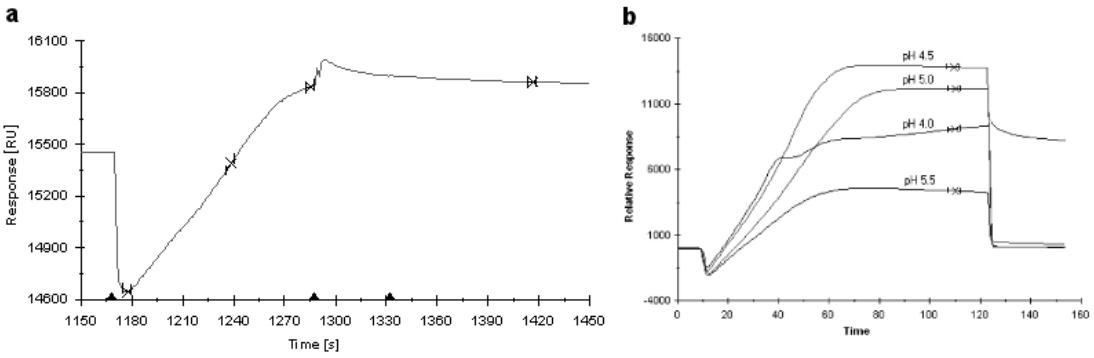

Figure 4.11: Preconcentration of His-EGFR on a non activated CMDP chip - a. The receptor was desalted in 10 mM NaOAc, pH 5.0 before injection (1160-1290 s). The signal is not returning to basis level after injection: the receptor is probably sticking on the surface due to denaturation. Unreferenced signal. **b.** Example of preconcentration scouting results. Binding increases as the pH is reduced from 5.5 to 4.5. At pH 4.0, the sensorgram is irregular and bound material does not dissociate from the surface at the end of the injection, indicating that the protein is aggregating or denaturing. The optimum pH for this protein is 4.5.[62] *Part* ***b*** *reprinted from* Biacore - Sensor surface handbook *(Ref. 62), copyright 2005-2007 GE Healthcare Bio-Sciences AB.*

the receptor at the surface without need of acidic buffer. The surface can be activated with EDC/NHS both before or after protein injection. The unreacted NHS-esters are then deactivated by injection of ethanolamine.

The main advantage of this strategy is that Ni^{2+} is only needed during the coupling step: after deactivation it can be washed away with EDTA, avoiding that way the unspecific binding problem. Two additional advantages are that proteins are coupled in a directed way, since they are all binding to the surface through the His-tag, and that untagged impurities are disfavoured. This results in a more homogeneous surface and, possibly, simpler binding curves[75].

The coupling procedure was successful with ∼1100 RU His-EGFR coupled on a NiP chip (Fig. 4.12). The binding was stable and a stable baseline could be reached. Preliminary analysis with NTA running buffers* detected binding of strep-ARNO-Sec7 when injected over the EGFR surface (Fig. 4.13). The negative control (GST) did not bind. However, the curves were not fitting to a simple 1:1 binding model and showed clear biphasic behaviour.

Other buffer systems were tested, such as HBS-P† , which was a good buffer for

*NTA running buffer: 10 mM HEPES, pH 7.4, 300 mM NaCl, 50 µM EDTA, 0.005 % Tween-20; NTA dispenser buffer: 10 mM HEPES, pH 7.4, 300 mM NaCl, 3 mM EDTA, 0.005 % Tween-20.

†HBS-P: 10 mM HEPES, pH 7.4, 150 mM NaCl, 0.005 % Tween-20.

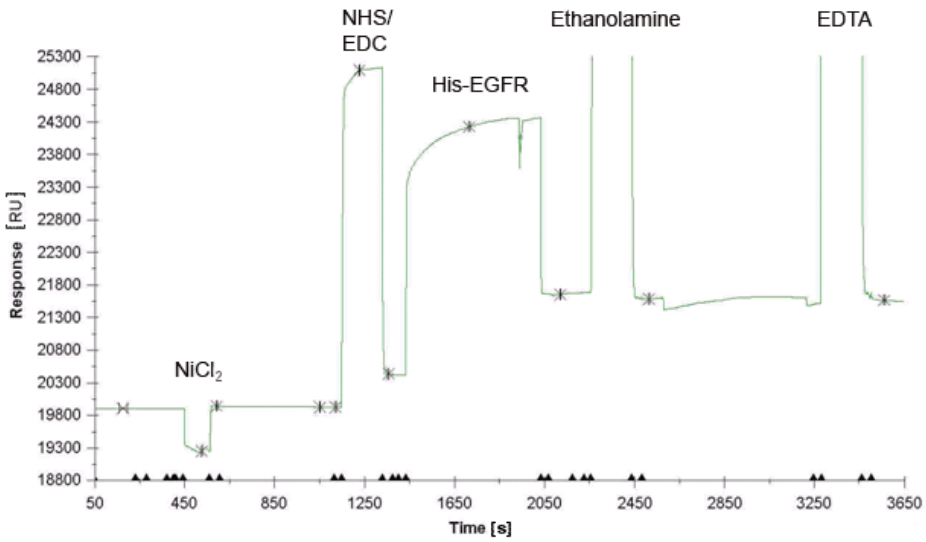

Figure 4.12: Covalent binding of His-EGFR on a NiP chip - The surface was first loaded with Ni^{2+}, activated with EDC/NHS and then the receptor was injected. Ethanolamine was used for deactivation and Ni^{2+} was washed away with EDTA. 1100 RU of receptor were immobilized. Unreferenced response.

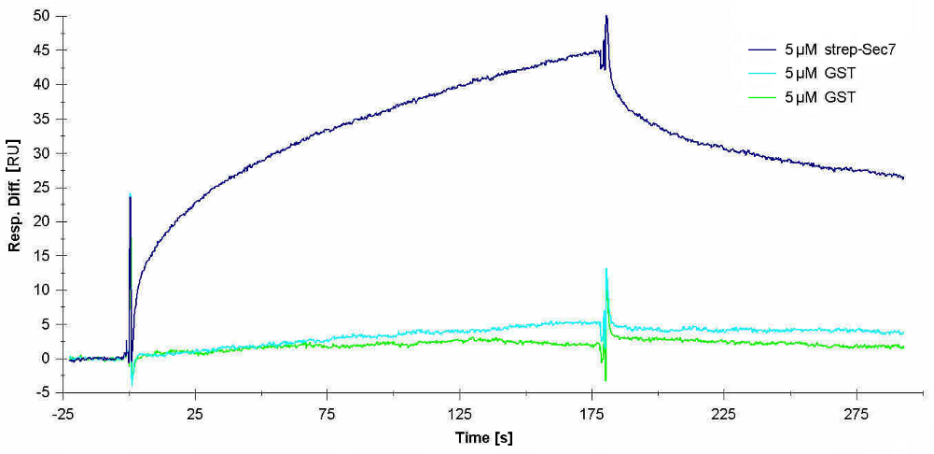

Figure 4.13: Comparison of the responses for ARNO and GST injected over His-EGFR covalently coupled on an NiP chip - Analysis performed in NTA running buffers. A clear signal is seen for ARNO but the shape of the curve is not simple exponential indicating a complex binding behavior. The control protein GST did not bind. Referenced with untreated Fc.

the receptor but not for ARNO. Indeed, strep-ARNO-Sec7 aggregated/precipitated in HBS-P during injection if diluted or during desalting in HBS-P. This was seen as jumping and bumpy curves in Biacore and extensive cleaning was needed after injection to get again sensible injection curves (data not shown).

In A-buffer* binding reached equilibrium and the dissociation phase was more sensible (Fig. 4.14). Unfortunately, EGFR was not stable in A-buffer and binding activity was completely lost after overnight incubation. Because of a delivery delay, no NTA/NiP chip was available for a new immobilisation, thus, meanwhile, I tried a coupling on CM5/CMDP chips without preconcentration.

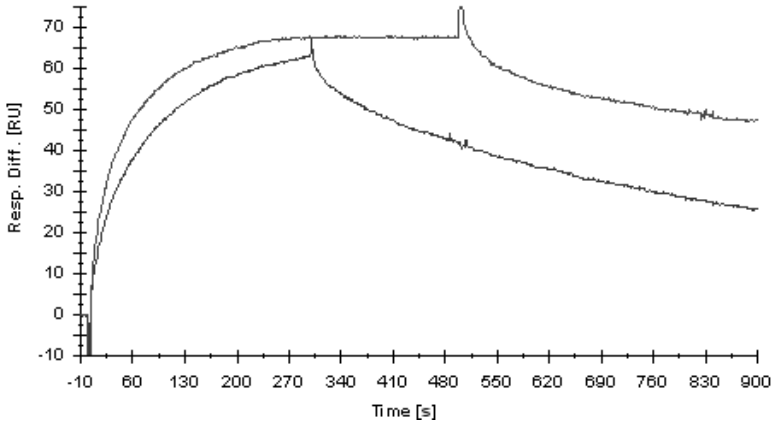

Figure 4.14: Responses in A-buffer reach equilibrium - Injection of 5 µM strep-ARNO-Sec7 in A-buffer over His-EGFR covalently coupled on an NiP chip was sufficient to reach equilibrium in ∼5 min. This buffer is unfortunately not suitable for the receptor which completely lost its binding activity by prolonged incubation. Referenced with untreated Fc.

4.3.1.4 Covalent coupling to carboxymethylated dextran chips without preconcentration

In principle, it could be possible to immobilise sufficient protein on a CM5/CMDP chip even without preconcentration and thus avoid the acidic conditions which are not tolerated by the receptor.

As first, I tested if it is possible to get enough ligand at the surface. In HBS-P buffer, the His-EGFR sticks on the chip after a simple injection without activation.

*A-buffer: 50 mM Tris, pH 7.8, 300 mM NaCl, 50 µM EDTA, 0.005 % Tween-20.

This is seen as an increase of the baseline level after injection (Fig. 4.15). Both chips from Xantec and Biacore were tested.

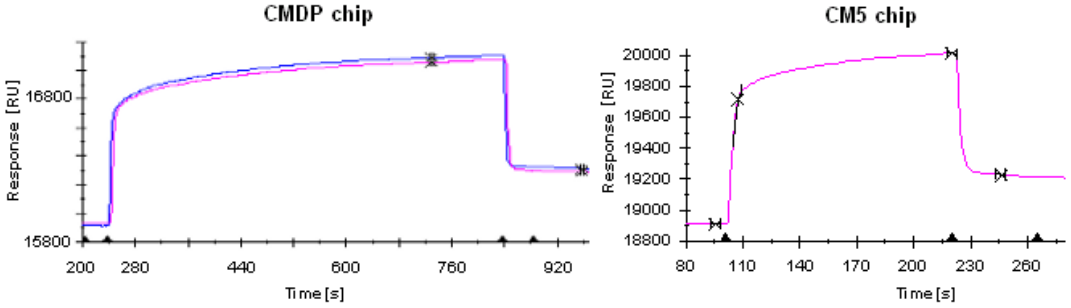

Figure 4.15: His-EGFR in HBS-P buffer sticks on carboxymethylated dextran surfaces - His-EGFR was injected over a non activated surface of a CMDP (left) or CM5 (right) chip (injection times: 250-840 s and 100-220 s, respectively). The receptor sticked on the surface and could only be removed under harsh conditions. Unreferenced responses.

Thus, a second EGFR construct (GST-EGFR) tagged with GST instead of the His-tag was tested. As GST-EGFR did not stick on the chip, it was used for coupling and ∼3400 RU receptor could be immobilized (Fig. 4.16).

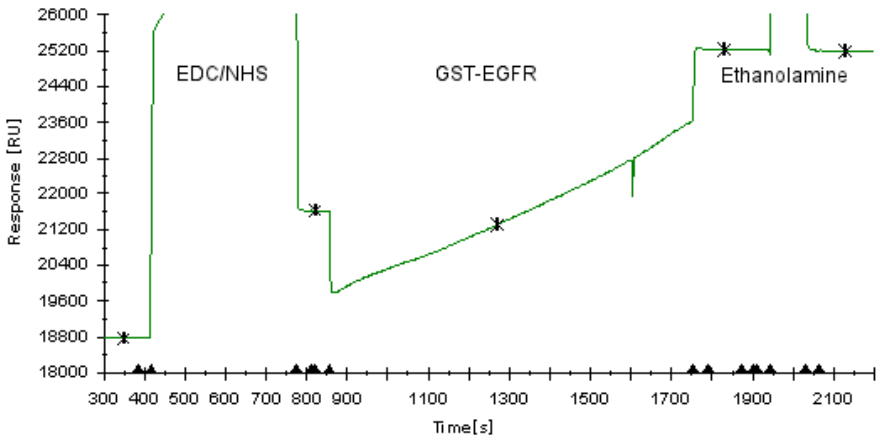

Figure 4.16: Coupling of GST-EGFR to a CM5 chip - The chip was activated with EDC/NHS solution (400-700 s) before injection GST-EGFR in HBS-P (850-1750 s) and deactivation with ethanolamine (1950-2050 s). 3400 RU EGFR were coupled. Unreferenced response.

Strep-ARNO-Sec7 bound in concentration dependent manner to immobilised GST-EGFR (Fig. 4.17) while the negative control (GST) did not bind. The shape of the curves showed some complexity (e.g., linearity at the end of the injection phase). Nevertheless, a fit to a 1:1 binding model was tried (black lines in Figure 4.17). Since

dissociation was slow, the acquisition time in the washing phase was too short to obtain a curvature of the dissociation response, a condition necessary for a precise determination of k_d. Longer acquisition times in the dissociation phase could not be achieved do to technical limitations. Thus, the binding parameters should be taken with reserve, since k_d values 2-3 times as high would still be compatible with the measured data and affect both k_a and K_D.

Some variation is present in the replicates in Figure 4.17. This is because the measurements shown were part of a data series where various regeneration conditions were tested. The most promising solution was 0.1 % SDS/1 % Triton X-100 in HBS-P, but still complete regeneration was not possible.

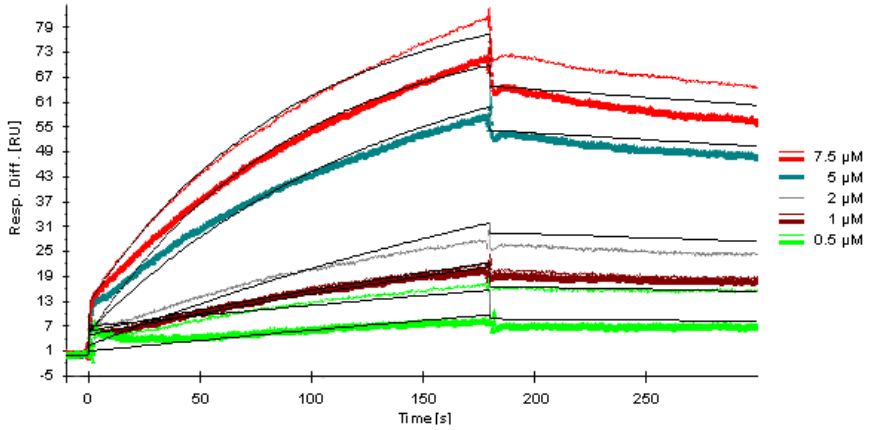

Figure 4.17: Binding of strep-ARNO-Sec7 to immobilised GST-EGFR - The responses are concentration dependent but display some complexity, as evident from the superimposed fit to a 1:1 binding model (shown in black). E.g., at the end of injection the responses become linear. Due to slow dissociation and limited acquisition time in the dissociation phase, k_d (and thus K_D) can not be determined precisely. The fit parameters (k_a = 1360 $M^{-1}s^{-1}$, $k_d = 6.2 \times 10^{-4}$ s^{-1}, K_D = 450 nM, R_{max} = 80 and χ^2 = 7.3) should therefore be taken with reserve. Responses referenced with untreated cell.

As a control, interaction of GST-MIG6, a known binding partner of the EGFR-ICD[44], and the receptor was analysed. The responses were concentration dependent and, except for some complexity at the start of the washing phase, overall well described by a simple binding model (Fig. 4.18). The K_D was estimated to 5 µM with k_a = 591 $M^{-1}s^{-1}$, $k_d = 3.0 \times 10^{-3}$ s^{-1}, R_{max} = 50 and $\chi^2 = 0.4$*. Again, longer acquisition of the dissociation phase would have allowed preciser determination of the binding

*Errors are not reported because the parameters results from a single global fitting operation.

parameters.

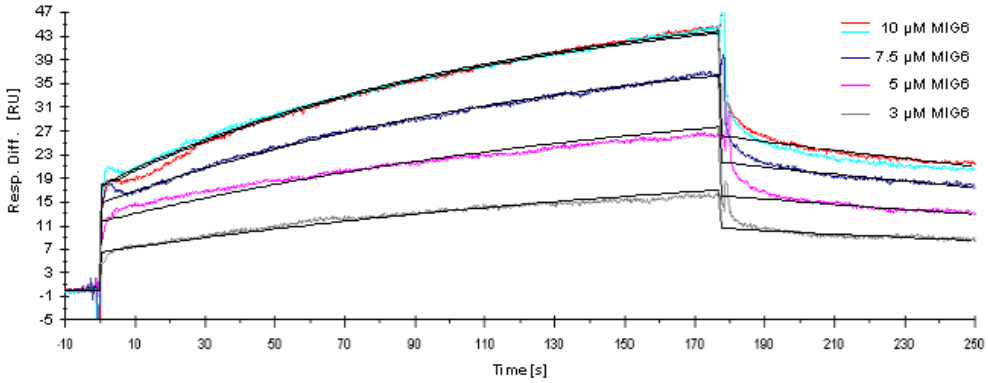

Figure 4.18: GST-MIG6 binding to immobilised GST-EGFR - The responses are concentration dependent and the association phase is very well described by a 1:1 binding model (black lines). Longer dissociation times would be needed to discern if the complexity an the start of the dissociation phase is relevant and allow preciser determination of k_{d}. Fit parameters: $k_{\mathrm{a}} = 591\ \mathrm{M}^{-1}\mathrm{s}^{-1}$, $k_{\mathrm{d}} = 3.0 \times 10^{-3}\ \mathrm{s}^{-1}$, $K_{\mathrm{D}} = 5\ \mu\mathrm{M}$, $\mathrm{R}_{\mathrm{max}} = 50$ and $\chi^2 = 0.4$. Responses referenced with untreated cell.

The analyte binding capacity gives the maximal response (in resonance units) one can expect from a surface saturated with analyte and is calculated as[62]

$$\text{analyte binding capacity (RU)} = \frac{M_{\mathrm{r}}(\text{analyte})}{M_{\mathrm{r}}(\text{ligand})} * \text{immobilised ligand (RU)}\,. \qquad (4.11)$$

In this case with the molecular weights M_{r}(GST-MIG6) = 32 kDa and M_{r}(GST-EGFR) = 88 kDa, and ∼3400 RU immobilized EGFR, the theoretical analyte binding capacity is ∼1240 RUs. However, the fit in Figure 4.18 estimates a maximal binding capacity ($\mathrm{R}_{\mathrm{max}}$) of 50 RUs which corresponds to only 4 % of the theoretical value. This probably means that only a minimal part of the coupled receptor was able to bind MIG6 and can have various explanations, as for example:

- some receptors loosed activity during the coupling, or binding was sterically hindered (this could be the case, e.g., if the receptor was coupled nearby the binding site)
- part of the receptors was already inactive in the stock solution
- the response from the coupling was not only given by the receptor but also by impurities (either in the ligand or coupling solutions).

In fact, the amount of coupled receptor was quite high for the conditions used (concentration, pH and coupling time). To check for a possible presence of impurities in the coupling solutions, a coupling in absence of protein was performed. Indeed, mock coupling of a reference surface lead to a significant increase in the baseline signal, i.e. the signal detected when flowing buffer (Fig. 4.19**a**). Up to then, coupling was performed with in-house reagents. Comparison with the original reagents from the 'Biacore coupling kit' showed that the last affected the baseline to a much lower extent (Fig. 4.19**b**). It is not clear what are the reasons for this difference, but it has to be taken into account when estimating the analyte binding capacity.

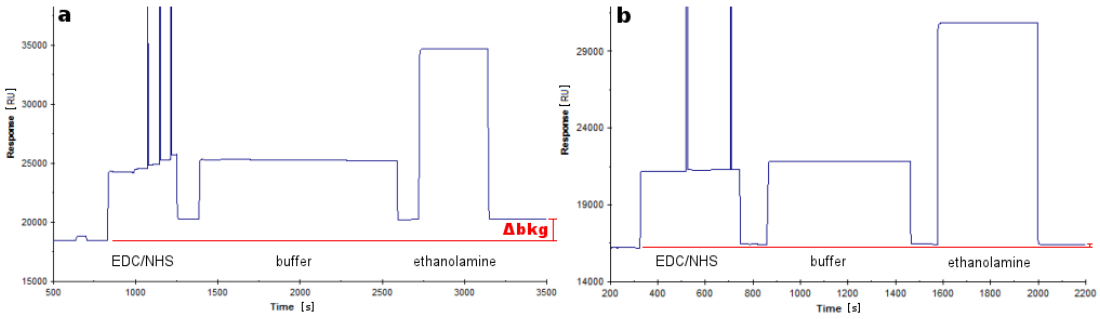

Figure 4.19: Mock coupling of the reference surface - The surface of a CM5 chip was activated with EDC/NHS solution, washed with coupling buffer, and deactivated with ethanolamine. **a.** When in-house coupling reagents were used, mock coupling led to a baseline increase of ∼1800 RU (Δbkg). **b.** Use of fresh 'Biacore coupling kit' increased the baseline of only ∼180 RU. *Reprinted from Ref. 76, p.42.*

Injection of strep-ARNO-Sec7 on the mock coupled surface (in-house reagents) lead to responses higher than on the GST-EGFR surface (Fig. 4.20). This, together with the increase in baseline, suggests that the coupling reagents used were possibly contaminated and could lead to false positive signals. The chips used in Section 4.3.1.3 (*Covalent coupling to nitrilotriacetic acid derivatised chips*) were treated with the same reagents and the results presented there are therefore also put in question. To confirm the binding results obtained, the experiments described above should thus be reproduced with the new coupling reagents.

4.3.1.5 Discussion

The experiments in Section 4.3.1 focused on Biacore analysis of a possible ARNO/EGFR interaction. Four different approaches for coupling of the receptor are described, with

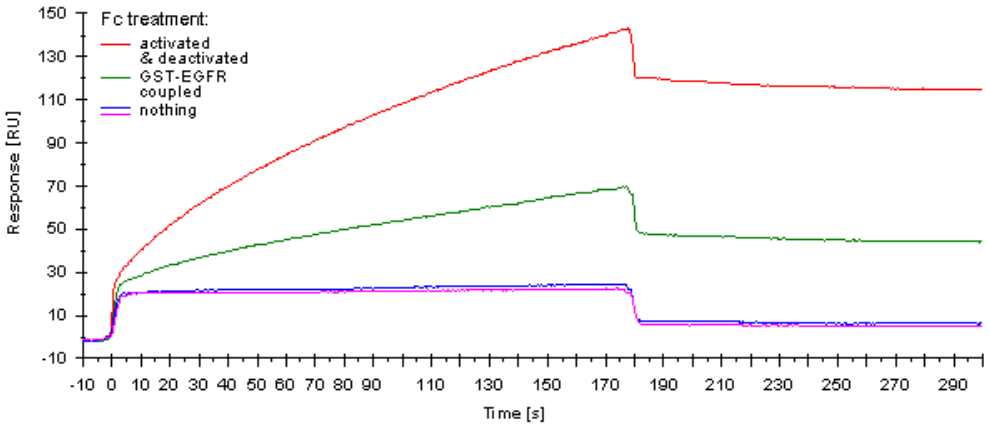

Figure 4.20: Unspecific binding of strep-ARNO-Sec7 to a mock coupled Fc - The surface of a CM5 chip was mock coupled with in-house reagents (see Fig. 4.19) The response of injection over the mock activated Fc (red) is higher than in the GST-EGFR coupled Fc (green). No binding was detected on untreated surfaces (blue and magenta).

special emphasis on common problems encountered and viable solutions. Here, only the results directly regarding the question if ARNO does bind the EGFR are discussed.

The first strategy was to immobilise His-tagged receptor on an NTA-derivatised chip surface via Ni^{2+} chelation. Since the high extent on metal-dependent, nonspecific protein adsorption to the biosensor surface made analysis very challenging, the method was discarded. Nevertheless qualitative results could already suggest an interaction.

The most promising method benefited of the enrichment of His-EGFR on an NTA surface to covalently couple the receptor with amine-coupling chemistry. This strategy qualitatively showed that ARNO and EGFR can interact. The complexity in binding, which prevented a quantitative analysis, can probably be reduced by the use of other buffers and new coupling reagents. Indeed, the screening of buffers already led to some improvement of the curve shapes. Additionally, using lower concentrations of ARNO could also significantly enhance the quality of analysis. Indeed, in Section 4.3.2.3 we show that concentrations of ARNO above 1 µM leads to complex responses independently of the coupled molecule. By now, we characterised the ARNO/EGFR interaction by fluorescence polarisation with a K_D of ~1 µM[50]. It could therefore be difficult to precisely determine the binding parameters with ARNO concentrations below 1 µM. In case of interest in the binding rates of the interaction, one should therefore consider the option of coupling ARNO instead of the receptor.

Despite of some complexity in the responses and limited dissociation data availability, covalent coupling of GST-EGFR to carboxymethylated dextran chips allowed to roughly estimate the binding parameters of ARNO and EGFR. Since dissociation could only be determined very imprecisely, K_D values 2-3 times as high as the obtained value of ~450 nM would still be compatible with the measured data and match well with the dissociation constant determined by fluorescence polarisation. Analysis of the interaction of GST-EGFR with the known binding partner GST-MIG6 gave a K_D of ~5 µM, which lies in the reported range[44,50].

For improvement of data quality, the same considerations as above are valid. The finding of a possible contamination of the coupling reagents inducing unspecific binding, could not be investigated in this project. Thus, the results described here should be reproduced with new coupling reagents.

4.3.2 Analysis of ARNO/inhibitor binding affinity

In the second part of this study, we were interested in establishing a method to analyse the binding affinity between the protein ARNO and its small organic molecule inhibitors by surface plasmon resonance (SPR). Since the discovery of SecinH3 some years ago[17] various derivatives were synthesised or identified by *in silico* screening, raising the need for a method for the analysis of their binding properties on cytohesins.

4.3.2.1 The derivatives of SecinH3

With the goal to identify improved cytohesins inhibitors, the group of Prof. Bajorath performed a virtual screening of chemical libraries formatted *in silico* and our group was involved in the *in vitro* analysis of the identified compounds.

Starting from limited information out of the primary hits from the HTS screening which lead to the identification of SecinH3[17,33] and the subsequent structure-activity-relationship (SAR) analysis, the virtual screening combined fingerprint similarity searching and support vector machine modelling[34]. Activity analysis *in vitro* of the 145 candidate compounds suggested by the virtual screening, identified 40 new structurally diverse inhibitors. Of these, 26 compounds were more active than SecinH3 in at least one of the three different biological assays tested[34]. Secin16 and Secin132 (Fig. 4.21) were the only two compounds with improved inhibition in all three assays[34].

Since Secin16 had a better IC_{50} in the nucleotide exchange inhibition ($3.1 \pm 0.5\,\mu M$ vs. $8.0 \pm 0.1\,\mu M$) we chose it for our SPR analysis.

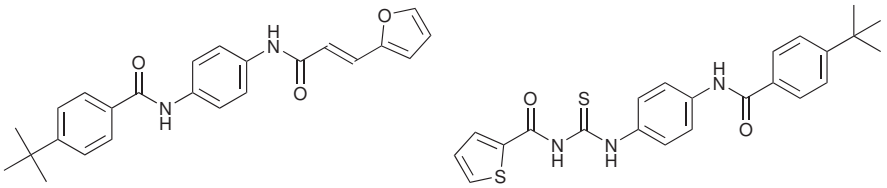

Figure 4.21: Structures of the virtual screening hit compounds - Secin16 and Secin132 were more active than SecinH3 in all three biological assays[34].

4.3.2.2 Measurements with immobilised protein*

Since for SPR measurements one of the interacting partners has to be immobilised on the chip surface, we had two possible experimental setups: either coupling the protein or the compound. Both methods should give similar results and have advantages and drawbacks.

On the one hand, coupling of the protein allows to measure several compounds on the same surface and does not require their derivatisation. Then again, one has to take into account the risk of loss of protein activity during the coupling process and of protein deterioration in the course of the experiments. Additionally, the detection of very low responses, as expected for this strategy because of the small dimensions of the compounds, is posing higher requirements for the quality of the reagents and the maintenance of the instrument.

On the other hand, coupling of small organic molecules usually requires their derivatisation and thus implies a considerable synthetic effort to afford a compound which retains its activity. All the same, a compound coupled surface offers the advantage of improved stability. As harsher conditions can be used, regeneration of the surface is generally more effective and loss of binding capacity is less frequently a problem. Moreover, binding of protein is inducing a high change in surface density leading to higher signal to noise ratios and easily detectable binding signals.

*The experiments in this and the next section (Sections 4.3.2.2 and 4.3.2.3) were performed by my master student Esteban Gutierrez.

There are various strategies for coupling a protein to an SPR-chip, including non covalent methods as well as covalent coupling with different chemistries. For our purpose it was important to have a stable surface, since any drifting of the baseline (as it occurs when protein dissociate from the chip) would additionally complicate our task of detecting small responses. We therefore directly excluded non covalent methods. Capturing biotinylated protein on streptavidine coupled chips is a non covalent approach which offers, thanks to the high affinity of biotin for streptavidine, surface stability similar to the covalent systems and has the advantage of being almost universally applicable. We still did not consider this approach as our first choice for two main reasons: first, the biotinylation of ARNO would have added an extra step to the procedure with additional risk of denaturation or loss of activity and, second, the streptavidine layer on the surface would have lead to an higher distance between the surface and the binding compound and reduced the sensitivity of the system further.

For these reasons, we decided to use EDC/NHS chemistry to covalently attach our protein to a CM5 chip. The optimisation of the coupling conditions and buffer system, as well as the measurements with various small molecules are described in detail elsewhere[76] and will not be reported here. Intensive control measurements showed that this setup was not suitable for the analysis of protein-compound interaction and we therefore went further to the second immobilisation approach.

4.3.2.3 Measurements with immobilised compound

To immobilise the compounds, we decided again to use NHS chemistry on CM5 chips.

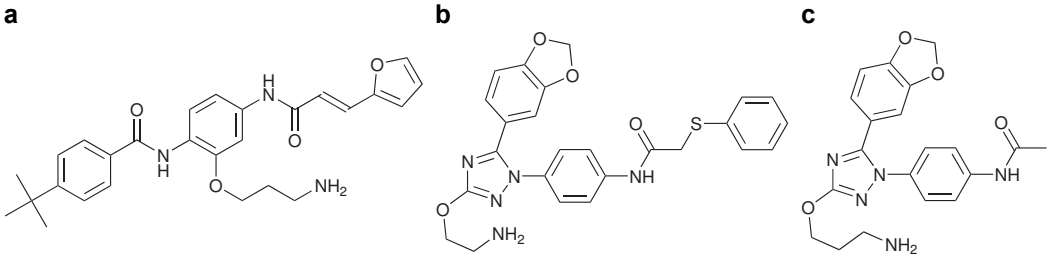

Figure 4.22: Derivatised Secin16 and control compounds - An amine derivatisation was needed to couple the compounds to the CM5-chip. Secin16 (**a**) and the negative control compound XH1009 (**c**) were synthesised with a propanolamine linker. The positive control compound SecinH3 (**b**) was already available with an ethanolamine linker.

This approach offers the advantage of a stable surface, but the compounds need to be

derivatised and the ideal linker length has to be determined experimentally. Secin16 and the negative control compound XH1009 were synthesised with a propanolamine linker, while an ethanolamine derivatised SecinH3 was already available and was coupled as positive control (Fig. 4.22).

Binding affinity of ARNO-Sec7 for Secin16

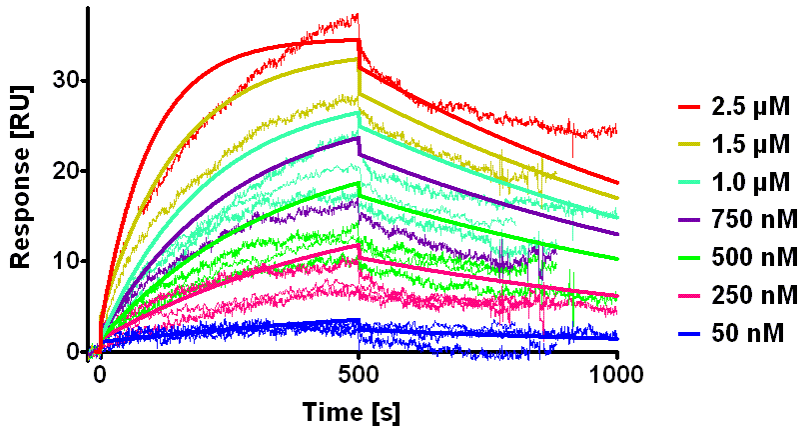

Figure 4.23: ARNO-Sec7 binding to immobilised Secin16 - Secin16 was covalently coupled to a CM5-chip and ARNO-sec7 at the indicated concentrations was allowed to flow on it. Four distinct data sets (three of them are shown) were fitted independently to a 1:1 binding mechanism and the average parameters used to simulate the expected binding curves (continuous lines). Although the data were of good quality and showed good reproducibility, the simulated curves are only poorly describing the measured ones.

The sensorgrams obtained by flowing ARNO-Sec7 on immobilised Secin16 were of good quality and acceptable reproducibility. Additionally, a clear concentration dependence of the responses was detectable. However, global analysis of the data was not satisfactory. Each experiment (with 4-6 protein concentrations) was analysed separately by global fitting and the mean of the kinetic parameters was taken to simulate the expected curves. As visible in Figure 4.23, the simulated curves does not describe adequately the experimental ones. A closer look to the curves shapes reveals that two different binding behaviours are present. In fact, the curves at low concentrations (till ∼500 nM) are exponential and reach equilibrium before the end of the injection phase (as expected from a 1:1 binding mechanism), while the ones at higher concentrations

are not reaching equilibrium any more but instead the response increase linearly at the end of the injection.

This is particularly evident when looking at Figure 4.24. In **a** the low concentration curves of the same experiments as in Figure 4.23 are shown. In **b** the results of measurements on a newly coupled chip are displayed. The four independent experiments were again fitted separately and the mean of the parameters used to simulate the curves in black. In this case, the simulation describes adequately the measurements. Moreover, the experiments on the different surfaces are described equally well. Since on the second chip less compound was coupled, the absolute responses are not comparable and this was taken into account for the simulation. The fact, that with different conditions (the surface density on the two chips is different), the same parameters are obtained confirms the quality of the data.

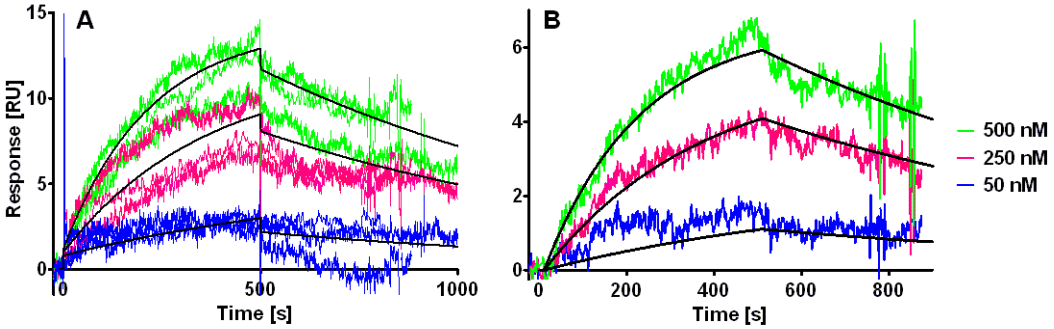

Figure 4.24: Analysis of ARNO-Sec7 binding to immobilised Secin16 at low protein concentrations - Out of the measurements shown in Figure 4.23 only the data with maximal protein concentration of 500 nM were taken for analysis. The measurements were performed at different days **(a)** and on distinct chips **(a, b)** with the indicated protein concentrations. The dissimilar absolute response between **a** and **b** is due to different compound densities on the two chips. The sensorgrams show good reproducibility and reasonable curves shape. Data analysis was performed as in the previous figure. The simulated curves are shown in black and describe the measurements adequately. The common binding parameters are $k_a = (7 \pm 1) \times 10^3$ $M^{-1}s^{-1}$, $k_d = (10 \pm 2) \times 10^{-4}$ s^{-1}, $K_D = 155 \pm 53$ nM, n = 4. Data are given as mean ± SEM.

To better display the complex curve behaviour, these parameters were used to simulate the expected response at high concentrations. As evident in Figure 4.25, the responses at higher concentrations are surprisingly higher than expected. This and the absence of a plateau of the curves suggest the presence of a second binding

mechanism, which take place at high protein concentrations. It is thinkable that at concentration above 500 nM, the Sec7 domain starts dimerising, leading to different measuring conditions. Additionally aggregation on the surface or on already bound protein could take place and be responsible for the non specific binding behaviour.

An additional problem of the high concentration curves is visible in the dissociation phase. In fact, the dissociation is not exponential, as expected for a 1:1 binding model, but definitely biphasic: the dissociation, rapid in the first seconds, is quickly slowing down and never complete. This is an additional reason to omit these data from evaluation.

When taking into account only the low concentration measurements a K_D of 155 $\pm$ 53 nM was determined ($k_a = (7 \pm 1) \times 10^3$ $M^{-1}s^{-1}$, $k_d = (10 \pm 2) \times 10^{-4}$ s^{-1}, n = 4).

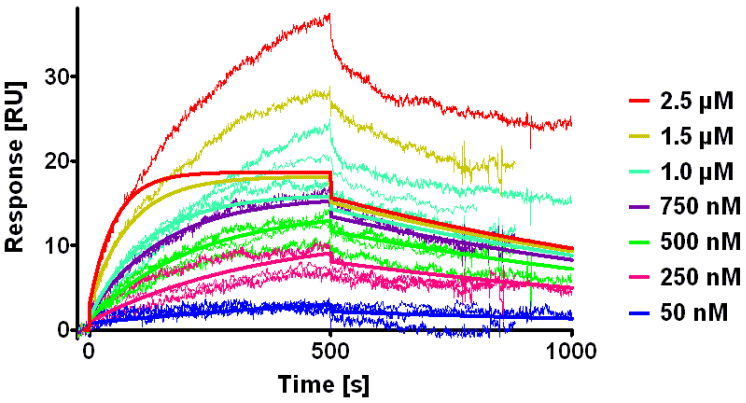

Figure 4.25: High concentrations of ARNO-Sec7 result in complex interaction with immobilised Secin16 - The data in Figure 4.23 are presented superimposed with the simulated curves of Figure 4.24 (continuous). The difference between simulated and acquired data, reveals that at higher protein concentrations the interaction does not behave as expected for a simple 1:1 binding mechanism. No equilibrium is reached, the dissociation is only partial and binding does not saturate.

Measurements with the positive control compound SecinH3

Binding of SecinH3 to ARNO had already been shown with other methods[17,77]. Thus SecinH3 was an interesting positive control. Unfortunately, no sensible sensorgrams were obtained on the SecinH3 surface (data not shown). Since already the buffer

injections behaved unpredictably, it is more probable that the surface was not ideal for SPR measurements, than this being caused by undesired interaction with the protein.

A possible cause of the problem, could be the different linker used for SecinH3 coupling. Since not available propanolamine derivatised, an ethanolamine derivatised SecinH3 was coupled to the chip. This shorter linker could have impaired the properties of the compound, leading to the strange response curves. A second linker was also tested and in fact, even if still not good, the sensorgrams showed a different behaviour.

A second possibility is that the immobilisation, and therefore concentration, of SecinH3 on the surface, can have lead to aggregation or other undesired interactions between the molecules. This would change the surface properties and perhaps lead to irregular responses. Because of this unexpected effects, the interaction between SecinH3 and ARNO could not be measured and quantified.

Measurements with the negative control compound XH1009

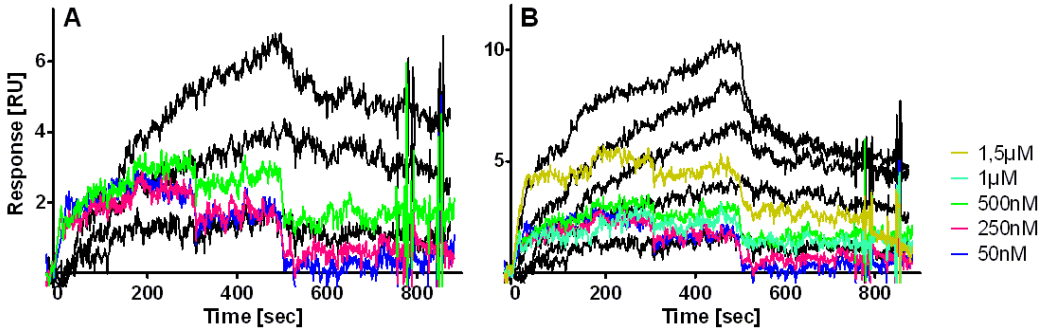

Figure 4.26: ARNO-Sec7 does not bind to the negative control compound XH1009 - The negative control compound XH1009 was coupled on the same chip as in Figure 4.24**b** and the sensorgrams acquired in parallel. The measurements on XH1009 are shown in color, the sensorgrams for binding to Secin16 were measured at the same concentrations and are shown in black for reference. ARNO-Sec7 does not exhibit significant binding to XH1009. The complex binding behaviour at high concentrations is visible in this figure too, in form of a sudden increase of response at the highest concentration (**b**, the same experiments as in **a** with all measured concentrations shown).

The compound XH1009, which has the same core structure as SecinH3 but has no inhibitory activity and does not bind ARNO[77] was used as negative control. ARNO-Sec7 did not show any significant binding to XH1009 (Fig. 4.26**a**). In Fig. 4.26**b** is

evident that at higher concentration (1.5 µM, yellow) the Sec7 behaves differently on this surface, too.

Measurements with a negative control protein

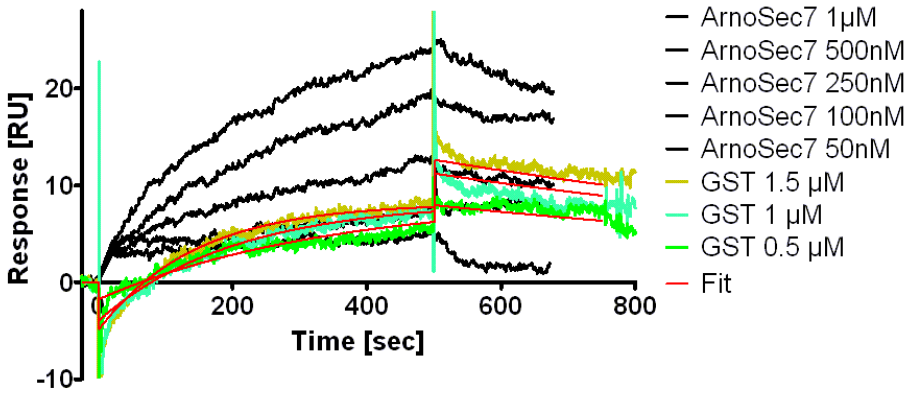

Figure 4.27: The negative control protein GST shows unspecific binding - GST was injected over covalently coupled Secin16. The sensorgrams for GST (in color) are displayed superimposed to the ARNO-Sec7 measurements (black). The absolute response is lower as for ARNO-Sec7, but the K_D (220 nM) is comparable. GST proved not to be an appropriate negative control, as discussed in the text (Section 4.3.2.4).

Having seen that ARNO-Sec7 is not binding to XH1009, we tested a negative control protein. For this scope glutathione S-transferase (GST) was flowed over the Secin16 surface. The results obtained were contradictory: while the sensorgrams shown in Figure 4.27 show binding of GST to Secin16, other data sets show no binding (data not shown). The sensorgrams displayed were fitted to a 1:1 binding model and, although some deviation are readily visible, the fit seems to almost adequately describe the curves. Though the absolute response is lower, the K_D obtained, 220 nM, is in the same range as that obtained for ARNO-Sec7. It has to be noted that for this measurements the concentrations of GST were higher than the acceptable ones for ARNO, while the sensorgrams showing no binding were measured at lower concentrations. It is questionable if GST is an adequate control, as discussed in detail in Section 4.3.2.4.

4.3.2.4 Discussion

Two different approaches were tested to evaluate the binding of small organic inhibitors to the Sec7 domain of ARNO by surface plasmon resonance. The first one involved

the covalent coupling of the protein to the chip surface and did not lead to biologically significant data. For the second one the amine derivatised compounds were immobilised and the protein flown on the surface. This setup led to sensible data but some artefacts were detected at high protein concentrations.

High ARNO concentrations alter binding

ARNO-Sec7 shows complex binding behaviour at high concentrations Measurements at different time points and on different surfaces showed a good reproducibility of the data but at the same time made evident that at least two binding modes were observed. The first one, seen at protein concentrations till ∼500 nM, has a standard 1 : 1 binding mechanism with exponential binding and dissociation curves and equilibrium at the end of the injection (Fig. 4.24). At higher concentrations unspecific binding seems to get the upper hand: the binding curves increase linearly at the end of injection and the dissociation curves becomes biphasic (Fig. 4.25). This effect is not surprising, since complex behaviour of ARNO at high concentrations was also observed with other methods, such as thermophoresis (A. Schmitz, unpublished data). This could be due to a solubility issue which leads to aggregation and unspecific interaction of ARNO either already in solution or on the surface (where the local concentration is even higher). Moreover, such artefacts seems to be common in SPR analysis as explained by Rich and Myszka: *Sometimes, the response continues to increase (and may become more complex) as the analyte concentration increases. Most likely this is due to heterogeneity in the ligand, analyte, or both. Weakly binding material or higher levels of non-specific binding is often observed at higher analyte concentrations.*[65]. In their annual survey of optical biosensor literature, they stress the importance of using adequate analyte concentrations to avoid these artefacts[65].

Secin16 interacts with ARNO-Sec7

Analysis of the measurement series at low concentrations reveals a dissociation constant of 155 ± 53 nM for ARNO-Sec7 and Secin16, a value which lies in a reasonable biological range. It is not possible to directly compare K_D and half maximal inhibitory concentration (IC_{50}), amongst other because of the dependence of the IC_{50} from the experimental conditions. Nevertheless, we can point out that this K_D could account for the IC_{50} of 3.1 µM observed in the nucleotide exchange assay[34].

Stumpfe *et al.* reported a K_D of 5 and 7 µM for the Sec7/Secin16 interaction determined by SPR and microscale thermophoresis respectively[34]. However his results can not be directly compared since different measuring approaches were used. Indeed, for their SPR measurements Stumpfe *et al.* immobilised the protein instead of the compound and used a special sensor chip with 3 D immobilization matrices (hydrogel surface). Moreover the Sec7 domain of cytohesin-1 and not ARNO was used[34]. For microscale thermophoresis ARNO-sec7 was used. However, there is no literature so far showing comparison of K_Ds determined by SPR and thermophoresis. It is thus difficult to determine how far the measured K_Ds are influenced by the measuring approach and the actual K_D probably lies somewhere between the values reported here and in Ref. 34.

Analysis of the absolute response

A surprising feature of these SPR measurements is the absolute response. For the first chip, the amount of immobilised ligand was estimated as 2300 RUs. Inserting the molecular masses of ligand (∼400 Da) and analyte (27 kDa), in Equation 4.11 gives a theoretical maximal response of over 150 000 RU. The theoretical value can strictly only be reached at infinite analyte concentration and contact time and the experimental value is also dependent on other factors, such as the activity of the ligand and the kinetics of analyte binding[62]. Nevertheless, the estimated maximal response of ∼17 RUs (Fig. 4.25) is suspiciously low.

Still, the binding response is dependent on the amount of compound immobilised, as depicted in Figure 4.24. In fact, the amount of compound coupled in **a** (2300 RUs) and **b** (800 RUs), correlates pretty well with the respective experimental maximal binding. Additionally, the equation above is actually only described for immobilised proteins and it is not clear if the same relation is valid for small organic molecules. Because of the quite different properties of these two type of molecule it would be reasonable if a correction term is needed. In particular, the maximal response could be limited by steric effects if the surface is densely modified with compound. Thus, although this irregularity should be kept in mind, we believe that this is not enough to discredit the data.

SecinH3 sensorgrams are irregular

Unfortunately, measurements with the positive control compound SecinH3 were not possible since no modified molecule with the right linker was available (the measurements with SecinH3 with shorter linker lead to non evaluable data, see Section 4.3.2.3 for possible explanations). It would be interesting to test if the propanolamine derivatised SecinH3 allows detection of binding data of quality comparable to that of the Secin16 measurements.

In a previous work, the dissociation constant of SecinH3 and ARNO-Sec7 was measured by SPR and isothermal titration calorimetry (ITC) giving K_Ds of 748 ± 70 nM[78] and 250 ± 5 nM[17], respectively. Although, since the SPR measurements were performed with a different approach and because of the dissimilar principle of ITC (where both interacting molecules are in solution) it is not possible to directly transfer this values to our system. Therefore, comparison of the binding affinity of Secin16 and SecinH3 is not possible at this stage.

The negative control compound XH1009 does not bind the Sec7 domain

Because of the aggregation propensity of ARNO and the complex binding behaviour observed, it was particularly important to have a good negative control. To exclude unspecific binding of ARNO-Sec7, the interaction between ARNO-Sec7 and immobilised XH1009 was analysed. As requested, no binding was detected (Fig. 4.26**a**). Interestingly, at concentrations above 1 µM a sudden increase of response was detected (Fig. 4.26**b**). This supports the hypothesis of a complex behaviour of ARNO-Sec7 at high concentrations.

GST is not an adequate negative control

As a cross check we analysed the behaviour of the completely unrelated protein GST on the Secin16 surface, but the outcome of this experiment was not conclusive. In fact, two different data sets, gave quite different results. In the first experiment GST concentrations between 0.5 and 1.5 µM were used and, although the absolute response was lower as that obtained for ARNO-Sec7, kinetic evaluation resulted in a K_D of ∼200 nM (Fig. 4.27), which is comparable to the K_D of the Sec7 domain. However, measurements with lower concentrations, did not show binding.

With the benefit of hindsight, we can say that GST is not an ideal negative control. GST was chosen because it is unrelated but of similar size to the Sec7 domain and readily available. Unfortunately, in our experiments these advantages were overwhelmed by the drawbacks. One of these is the constitutive dimeric state of GST, which can cause complex behaviour of the protein on the chip. Moreover, GST has not only a broad specificity for its ligands but it is also reported to bind a broad spectrum of nonsubstrate ligands[79,80], a peculiarity for which it earned, at the time of its discovery, the name *ligandin*[81]. It is probably this kind of unspecific binding that we detected in our measurements, an observation corroborated by thermophoresis data where GST was shown to bind to all binding partner tested (A. Schmitz, unpublished). All together, it is now clear that the data acquired for GST can not be taken in consideration for our evaluation and that a better negative control protein should be selected. Ideally this protein should have similar size and isoelectric point as the Sec7 domain, be monomeric in solution and do not show noticeable unspecific binding behaviour. E. Gutierrez's current measurements of additional control proteins seems to confirm the specificity of the detected ARNO/Secin16 interaction.

5

Capture compound mass spectrometry

When a new active compound is identified, the question about its specificity for the target protein directly arises. Traditionally, this is investigated *in vitro* with binding or activity assays with more or less related proteins. Unfortunately, this process is not only extremely time consuming but often leads to biologically non relevant data. Moreover the number of proteins which can be tested is minimal compared to the proteome.

The interest on identifying possible side targets already at an early stage of compound characterisation, led to the development of new methods in the last years. This process was speeded up by the revival of whole-cells based screening for drug discovery, since the target protein of active small molecules identified that way are unknown and need to be identified in a successive stage[82].

A very innovative method which allows the one-shot test of an entire proteome is known as activity/affinity based protein profiling (ABPP) or capture compound mass spectrometry (CCMS). Goal of this project, was to establish CCMS for our inhibitors and SecinH3 was taken as model compound.

5.1 The concept of capture compound mass spectrometry (CCMS)

Activity based probes (ABPs) were first developed as a tool to collectively monitor the activity of enzymes with similar reaction mechanisms[83]. Their use was initially limited

to irreversible inhibitors which covalently bind to the target protein (see Ref. 84 and 85 for review) but the method was soon adapted also to non covalent inhibitors[86–88].

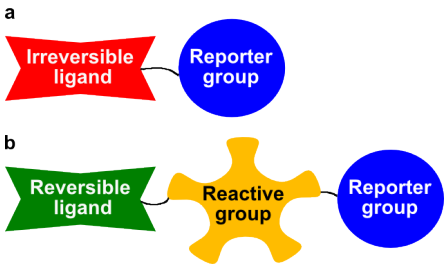

Figure 5.1: Activity/affinity based probes - a. Conventional activity based probe. The irreversible inhibitor acts as reactive group and is covalently bound to the enzyme. For detection and enrichment, reporter groups like biotin or fluorescent tags are used. **b.** Photoreactive affinity based probe. Upon irradiation, the reversible inhibitor is covalently bound to the interacting protein via the photoreactive group.

A conventional activity based probe is made up of three units: a reactive group, a linker, and a reporter group. In the case of an irreversible inhibitor, the reactive group is the ligand, which covalently binds to the enzyme (Fig. 5.1**a**). The reporter group can be a biotin or a fluorescent tag and is needed for the detection of the modified protein. Since the analysis is usually performed under denaturing conditions, e.g. gel electrophoresis, the method could not be applied to reversible inhibitors. Adding a photolabile unit to the probe, as first described by Hagenstein *et al.*[86], is an elegant solution to circumvent this limitation (Fig. 5.1**b**). This way, the inhibitor acts as a Trojan Horse by recruiting the probe to the protein which is then irreversibly modified by the photoreactive group upon irradiation[86–88].

An important step toward the development of a powerful method was the combination of ABPP with mass spectrometry (MS)[89–92]. This association allows, not only to readily identify the proteins modified by a probe, but also to determine their binding site. As shown in Figure 5.2, the whole process, called CCMS[93] when reversible ligands are used, consist of 6 steps. First, the probes are incubated with the proteome till binding equilibrium is achieved. Upon UV irradiation (**a**), the probes binds covalently the interacting proteins. A pull down with beads specific for the reporter group (**b**) allows enrichment of the labelled proteins. Then, a proteolytic cleavage is performed directly on the beads (**c**). Finally, the supernatant is analysed by liquid chromatography tandem mass spectrometry (LC-MS/MS) (**d**) and the data obtained is compared with a

protein database for identification (**e**). For the identification of the site of modification, the labelled peptides can be eluted from the beads and analysed the same way with LC-MS/MS.

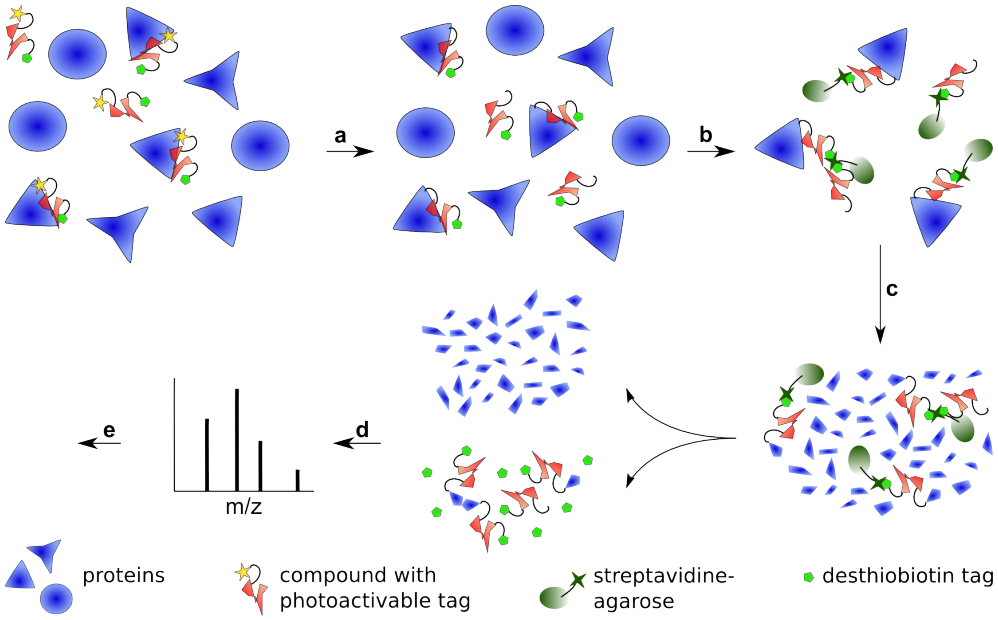

Figure 5.2: Capture compound mass spectrometry - Work flow of a standard CCMS experiment. The photoreactive probe is preincubated with the proteome. Upon UV irradiation (**a**), the probe reacts and binds covalently the interacting protein. Labelled proteins are enriched by pull down with streptavidine beads (**b**). Proteolytic cleavage is performed directly on the beads (**c**). The supernatant is analysed by LC-MS/MS (**d**) and the data obtained compared with a protein database for identification (**e**). Labelled peptides can be eluted from the beads for identification of the site of modification.

5.2 Mass spectrometry in proteomics

The first measurement of the mass-to-charge ratio (m/z) of various gaseous ionised atoms or molecules by Sir Joseph John Thomson signed the start of mass spectrometry in 1910. Though great advancements of the method in the subsequent years, for the first 80 years its application was essentially limited to small chemicals[94]. The breakthrough was made at the end of the Eighties with the development of the soft ionisation methods electrospray ionization (ESI)[95] and matrix-assisted laser desorption/ionization (MALDI)[96,97]. Allowing gentle ionisation of large non-volatile molecules, they extended MS to protein analysis. This, together with the quick growth of protein sequences

databases, thanks to the extensive genomics efforts, led to the rapid establishment of MS as a standard method for protein identification and analysis[98].

5.2.1 The working principle of a mass spectrometer

The three main components of a mass spectrometer are the ionization source, mass analyzer, and detector.

The ionisation source is needed to introduce the sample in the system. It is at this stage that the molecules are ionised and brought into (pseudo) gas phase. As MS can only measure charged molecules, this is an important step, which was precluded to proteins before the invention of MALDI and ESI. For MALDI[96,97], the sample is co-crystallised with a matrix which is then vaporised with UV laser pulses to liberate the peptides or proteins as gaseous ions. In ESI[95], the molecules are dissolved in liquid which is sprayed under high voltage to form small charged droplets. Evaporation of the solvents leaves singly and/or multi charged molecules which can enter the mass analyser.[98,99]

There are various types of mass analyser (e.g. time-of-flight (TOF), quadrupole, ion trap) which all rely on the application of electromagnetic fields to separate the ions according to their m/z.

The detector measures the number of ions having the m/z ratio corresponding to the applied fields. The display of ion intensity versus m/z is called a mass spectrum.

For this project, a mass spectrometer with ESI source and quadrupole ion trap analyser was used.

5.2.2 Peptide sequencing by tandem MS

Tandem MS allows the acquisition of additional information by performing two subsequent mass analyses. In the first step, a mass spectrum is acquired and "precursor ions" are selected for further processing. These ions are then fragmented and the resulting "product ions" are again separated according to their m/z and recorded as an MS/MS spectrum[98].

Fragmentation is usually achieved by collision induced dissociation (CID), in which the energy imparted by collisions with an inert gas (such as nitrogen, argon, or helium) causes the peptide to break apart[100]. At low collision energies (10-50 eV), the

product ions are primarily formed by cleavage at the peptide bonds[98]. The Roepstorff-Fohlmann-Biemann[101,102] nomenclature of the possible ions is shown in Figure 5.3. Ions are labelled as a, b, and c when the charge is retained by the amino-terminal fragment and x, y, and z if the charge is retained on the C-terminal side. In ion trap instruments y and b ions predominate[100].

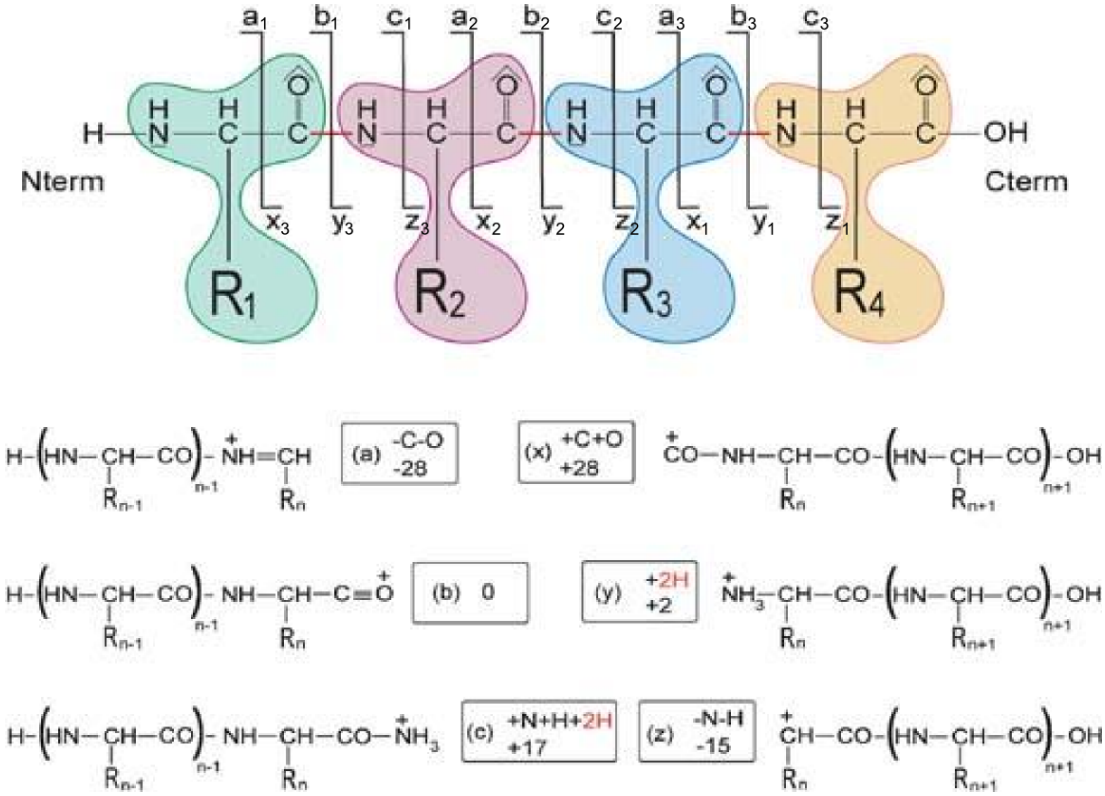

Figure 5.3: Product ions of collision induced dissociation - This figure shows different possibilities of fragmentation for a tetrapeptide (top) and the produced fragment structures (bottom). Ideal N-terminal fragments are computed by adding the mass of the N-terminal group (H) and all amino acid nominal masses before the cleavage position. Ideal C-terminal fragments are computed by adding the mass of the C-terminal group (OH) and all amino acid nominal masses after the cleavage position. Each ion type is characterized by an offset that represents the mass difference in Daltons between the observed mass and the corresponding N- or C-terminal ideal fragment. For example, the b-ion type offset is 0 (Da), because the mass of a b-ion type exactly corresponds to an ideal N- terminal fragment, whereas an a-ion type offset is -28 (Da), because a-ions lose a carbonyl and an oxygen atom compared to an ideal N-terminal fragment.[98] *Reprinted from Mass Spectrometry Reviews (Ref. 98), copyright 2006, with permission from John Wiley and Sons, Inc.*

As shown in Figure 5.4, in an ideal MS/MS spectrum each peptide fragment differs from the neighbour by exactly one amino acid. The mass difference between two adja-

cent peaks therefore theoretically allows to infer the amino acid sequence. In practice, missing peaks and noise can highly complicate the analysis and *de novo* sequencing rely on high quality data and an experienced eye. Indeed, even the leading *de novo* algorithms are still unreliable[100,103].

This problem was circumvented by the combination of database search with *de novo* sequencing[104], as described in the next section. This strategy takes advantage of the fact that for most peptides short partial sequences are easily determined. Using this "peptide sequence tags" for database matching allows rapid protein identification.

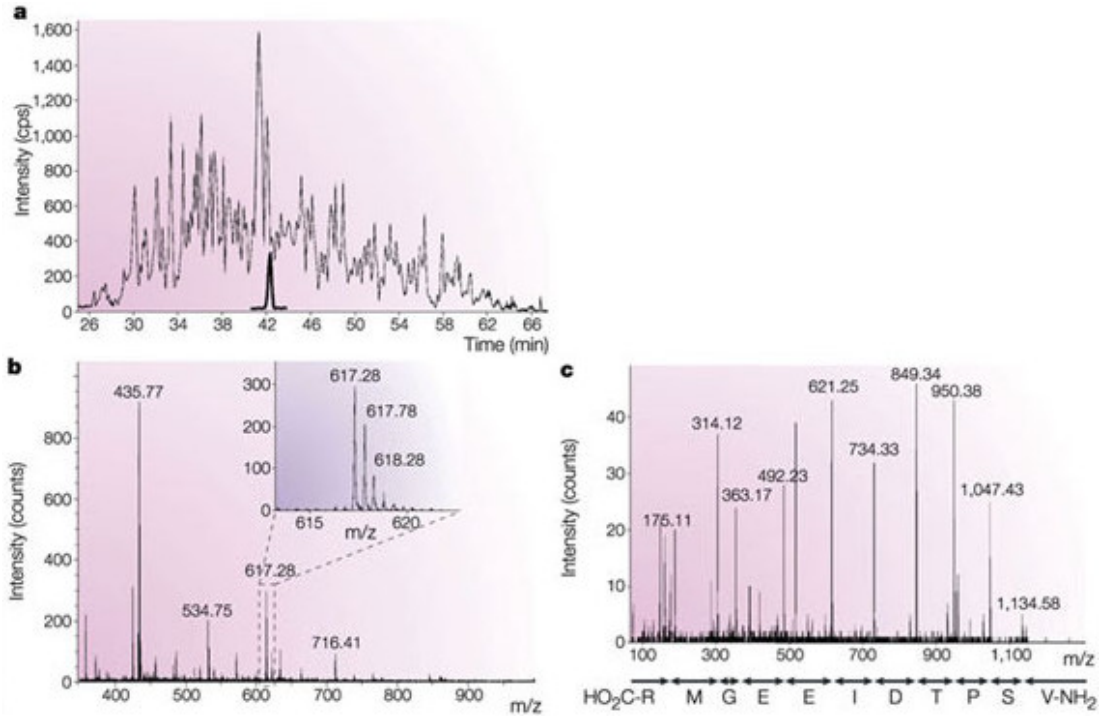

Figure 5.4: Mass-spectrometry traces - **a.** The total ion intensity from all the mass spectra that were recorded during the LC-MS run is shown as a function of elution time (total ion chromatogram). Shown in bold is the trace for the intensity of one particular ion, which elutes within a 40-second window approximately 42.5 minutes into the gradient (extracted ion chromatogram). The area under this curve represents the total signal of this peptide. **b.** The mass spectrum of the peptides that were eluted 42.4 minutes into the gradient. The insert shows the m/z values around the peptide ion of interest, which are indicative of the resolution and allow the charge state to be derived. **c.** The MS/MS spectrum of the peptide ion of interest (highlighted by a dashed box in part b). The mass differences between this y-ion series indicate the amino-acid series, which is shown below the spectrum. As this is a y-ion series, the sequence is written in the carboxy-to-amino-terminus direction going from left to right.[100] *Adapted by permission from Macmillan Publishers Ltd: Nature Reviews Molecular Cell Biology (Ref. 100), copyright 2004.*

5.2.3 Protein identification by database searching

The so called "top-down" strategy of protein analysis involves the ionisation of intact proteins and his fragmentation inside the spectrometer. Although this approach allows a more complete characterisation of protein isoforms and modifications[105], its use is still limited by the low throughput[106], expensive instrumental set up, lower sensisitvity and limited availability of specific bioinformatic tools[107].

Widespread is the "bottom-up" strategy, which analyses peptides generated by proteolytic digestion. Because of their smaller size and better solubility, peptides are readily separated by chromatography and can be analysed on every MALDI and ESI instrument[107]. Trypsin, a protease which cleaves at the carboxyl side of lysines or arginines[108], is routinely used in proteomic analysis because of the favorable length of the obtained peptides. Moreover, every tryptic peptide carries at least two protonation sites (the N-terminal amino group and the C-terminal lysine or arginine) for efficient ionisation and detection[109].

There are two main bottom-up approaches: peptide mass fingerprinting (PMF)[110–114] and tandem MS based data base search[104,115]. In PMF the peptide masses obtained experimentally are compared with the theoretical masses from *in silico* digest of a reference database. Since peptides with very similar masses can be derived from different proteins, this method requires an accurate mass determination, a large number of peptide matches and a high percentage of protein sequence coverage for a reliable protein identification. The second approach relies on the possibility of deriving sequence information from CID spectra, as described in the previous section. The simultaneous knowledge of peptide mass and (partial) sequence improves considerably the database search as it is theoretically possible to correctly identify a protein with a single MS/MS peptide match[109].

5.3 Results and discussion

Capture compound mass spectrometry can be applied to any binding molecule for which the necessary modifications are possible. This is particularly interesting for our group, where new inhibitors discovered by high throughput screening need to be characterised. The goal of this project was to establish a generally applicable CCMS protocol in our

group. Our best known small molecule inhibitor, SecinH3, was chosen as molecule of interest. The first step was thus the synthesis of a photoreactive SecinH3 probe.

5.3.1 Design and synthesis of a photoreactive SecinH3 probe

As described in Section 5.1 (p. 66), photoreactive affinity based probes are composed of three elements: an affinity unit, in our case SecinH3, responsible for the interaction with the target protein; a photoreactive moiety, which upon irradiation covalently binds the probe to the enzyme; and a reporter group, for detection and enrichment of the labelled protein.

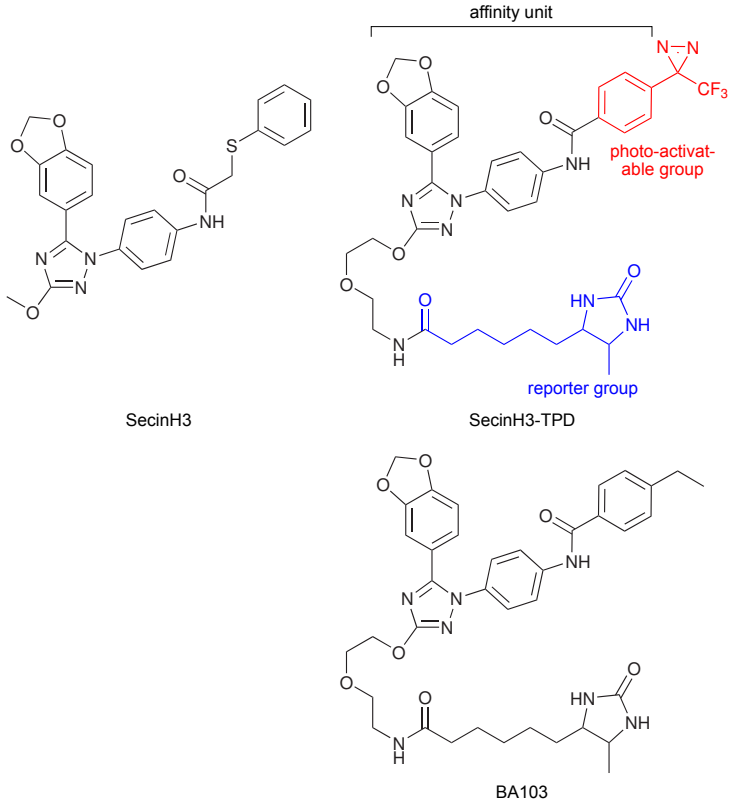

Figure 5.5: Structures of SecinH3, SecinH3-TPD and BA103 - In the photoreactive probe SecinH3-TPD, the thiophenyl group is substituted by the photo-activatable group trifluoromethylphenyl diazirine (red). The reporter group is desthiobiotin linked at the 3-position of the triazole ring (blue). The non photoreactive derivative BA103 serves as negative control.

An ideal photo-activatable group should be small, stable in the absence of light, as well as highly and indiscriminately reactive upon irradiation at wavelengths which do not damage the biological sample[116]. Commonly used photo-activatable groups include aryl azides, benzophenones, and 3-aryldiazirines. Though the advantages of trifluoromethylphenyl diazirine (TPD)[117] over the other photoactivatable groups were early

recognised, its use in photoaffinity labelling was limited in the past by synthetic restrictions and limited availability of TPD derivatives with convenient reactive groups[118,119]. The recent establishment of various diazirine compatible reactions[118,120] and the increasing commercial availability of TPD derivatives, now allows their more widespread use. The exceptional reactivity at relatively long wavelength (∼360 nm), the superior properties, and its structure make of TPD an ideal photo-activatable group for the modification of our small organic cytohesin inhibitor SecinH3. In particular the lack of formation of slowly reacting isoforms, as observed for benzophenones and aryl azides[116,121,122], is an important advantage of TPD which make it suitable for binding site identification experiments.

Structure-activity-relationship (SAR) studies on SecinH3 showed that the phenyl group is necessary for inhibition but its modification does not seriously affect activity and that the alkoxy substitutent is not essential (Ref. 77 and X. Bi, A. M. Hayallah and M. Famulok, unpublished results). These results offered the terminal thiophenyl group as an ideal target for effective substitution with the TPD moiety (Fig. 5.5) by amine coupling reaction with commercially available 4-(1-Azi-2,2,2-trifluoroethyl)benzoic acid (**TPD-COOH**).

Since extension of the methoxy group at the 3-position of the triazole ring is possible without significant loss of inhibitory activity[77], we decided to take it as anchoring point for the reporter unit. For a straightforward enrichment of labelled protein at later stage, we decided to use desthiobiotin, coupled through an aminoethoxyethanol spacer, as affinity element (Fig. 5.5). The biotin analogue desthiobiotin binds less tightly to streptavidin and can be easily displaced by biotin[123]. Thus, proteins labelled with desthiobiotin can be recovered after enrichment.

5.3.1.1 Retrosynthesis

The photoreactive building block **TPD-COOH** is not only quite expensive, but it requires also special care during handling (e.g. avoid exposition to light). Thus, we wanted to develop a synthetic strategy where the TPD moiety could be inserted efficiently in the last step.

Starting from the synthesis of the biotinylated SecinH3 precursor **Bio-SecinNH$_2$** described by X. Bi (Fig. 5.6, unpublished data), we formulated the retrosynthetic plan

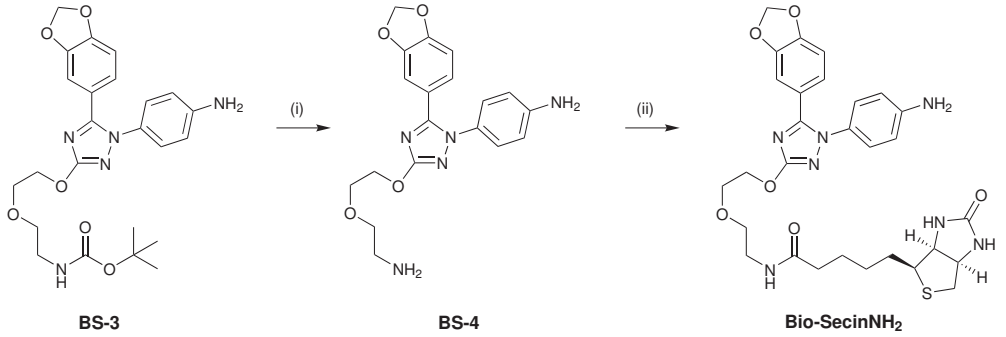

Figure 5.6: Biotin coupling - This procedure for the synthesis of the biotinylated SecinH3 precursor Bio-SecinNH_2 was reported By X. Bi (unpublished data): (i) dichloromethane, TFA, 1 h, RT (82 %); (ii) D(+)-biotin, HBTU, N-ethyldiisopropylamine, DMF, 30 min, RT (76 %).

in Figure 5.7 for the synthesis of SecinH3-TPD. Both **TPD-COOH** and desthiobiotin are coupled via peptide coupling reactions to **BS-4**. This can be obtained by deprotection of **BS-3**, which was available in sufficient amount.

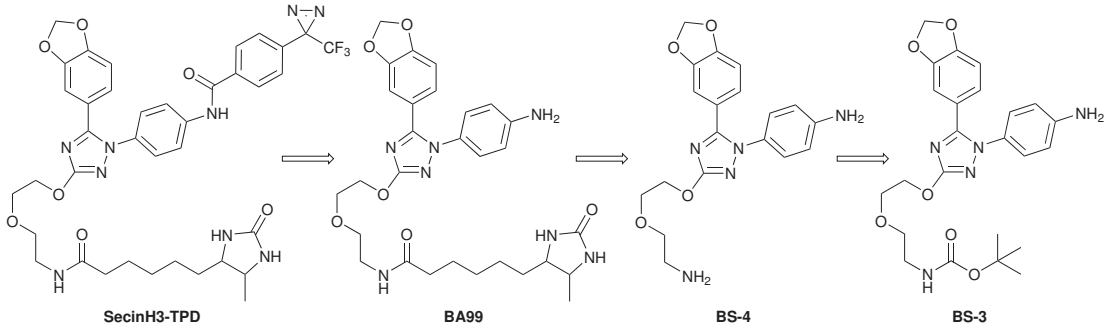

Figure 5.7: Retrosynthetic path for the synthesis of SecinH3-TPD - Peptide coupling is used to react TPD-COOH and desthiobiotin with **BS-4**, which is produced by deprotection of **BS-3**.

In fact, the reaction of **BS-4** with desthiobiotin (activated with HBTU and Hünig's base) did not give the desired product: probably the reaction was not selective enough. A new retrosynthetic analysis was therefore necessary. In the new retrosynthetic plan, **BA99** is obtained by hydrogenation of the nitrophenyl compound **BA96**. Desthiobiotin can selectively react with the free amine of **BA94**. Cyclisation of **BS-1** leads to the 1,2,4-substituted triazole ring in **BA91**.

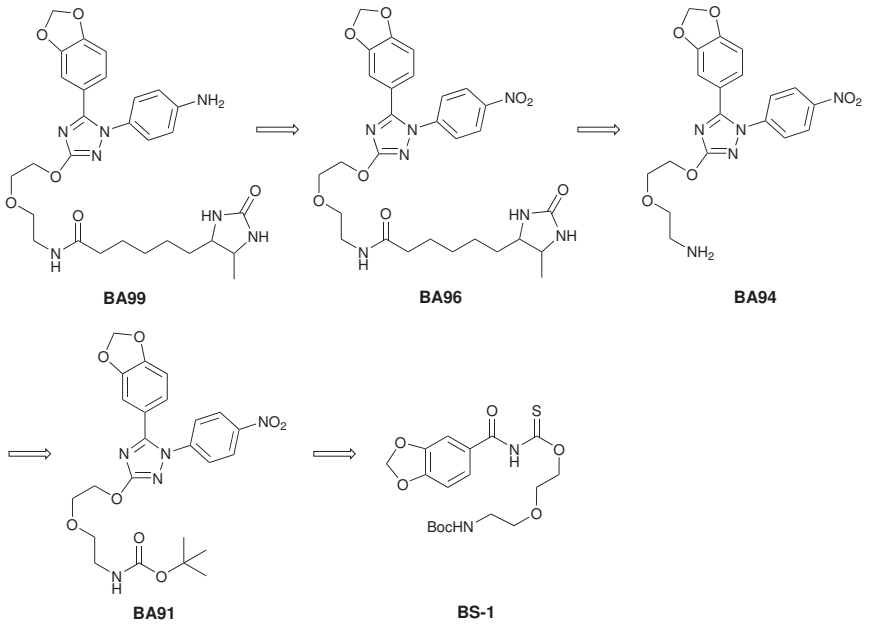

Figure 5.8: Retrosynthetic analysis for the synthesis of BA99 - Selective desthiobiotinylation can be achieved on **BA94**. The 1,2,4-substituted triazole derivative **BA91** is obtained by cyclisation of **BS-1**

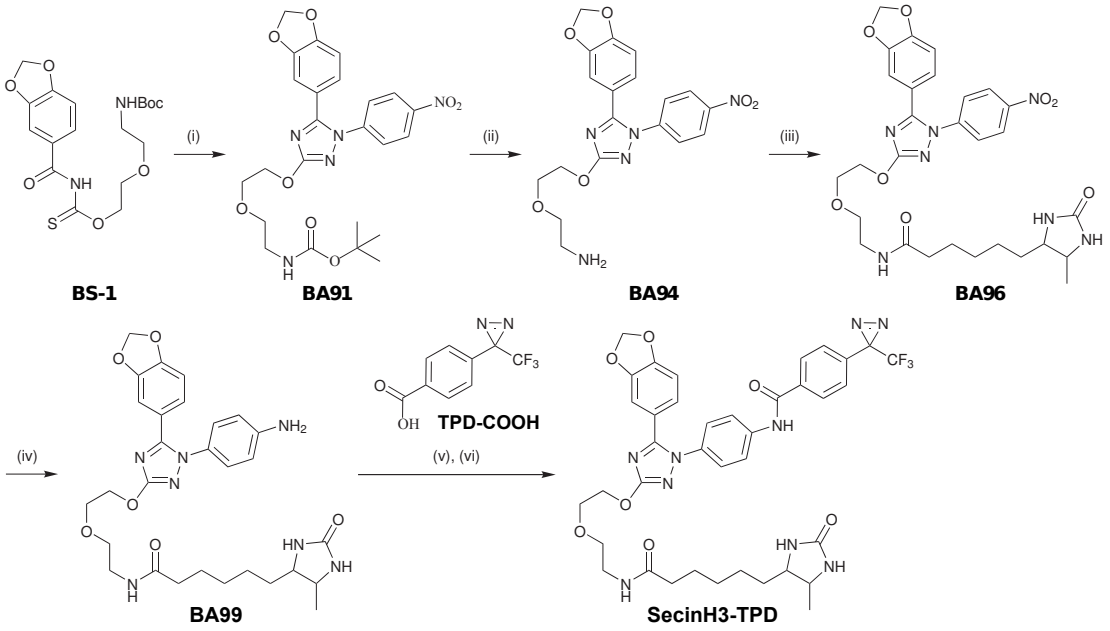

Figure 5.9: Synthesis of the photoreactive SecinH3 probe SecinH3-TPD - (i) Nitrophenyldiazirine, EtOH, 80°C, 8 h; (ii) Formic acid, 2.5 h; (iii) Desthiobiotin, HBTU, *N*-ethyldiisopropylamine, DMF, 1 h; (iv) 10 % Pd/C, EtOH/THF, 9 bar H_2, 2 h; (v) **TPD-COOH**, oxalylchloride, THF, DMF (catalytic), 2.5 h; (vi) **BA99**, *N*-ethyldiisopropylamine, CH_2Cl_2, 2 h.

5.3.1.2 Synthetic way

The synthesis of SecinH3-TPD is depicted in Figure 5.9. The 1,2,4-substituted triazole ring **BA91** was obtained by cyclisation of **BS-3** (Ref. 77) with 4-nitrophenylhydrazine. Deprotection of **BA91** with formic acid lead to **BA94**, which was then coupled to desthiobiotin by standard peptide coupling reaction. The desthiobiotinylated SecinH3 core **BA99** was obtained by palladium catalysed hydrogenation of **BA96**.

The final coupling was optimised in a model reaction on a non photoreactive benzoic acid: 4-ethylbenzoic acid was reacted to 4-ethyl-benzoyl chloride with oxalylchloride and then coupled to **BA96** to give the negative control compound **BA103** (Fig. 5.5).

Analogously, the benzoic acid **TPD-COOH** was converted to the benzoyl chloride before reaction with **BA99** to give the final product **SecinH3-TPD**. All steps gave the desired product with good purity and yield.

5.3.2 Determination of the labelling conditions *in vitro*

Having synthesised the photoreactive probe SecinH3-TPD, the ideal labelling conditions had to be established. The amount of protein labelled upon UV irradiation correlates directly with the portion of protein bound to the compound at equilibrium. Therefore, to define the initial working concentrations for both protein and compound, I first calculated the expected complex ratio at various concentrations.

5.3.2.1 Theoretical binding ratio at various compound and protein concentrations

Assuming a simple 1:1 binding behaviour, the amount x of protein:compound complex at equilibrium depends from the initial concentration of protein ([P]) and compound ([C]), and the dissociation constant (K_D) according to

$$x = \frac{[\mathrm{P}] + [\mathrm{C}] + K_\mathrm{D} - \sqrt{([\mathrm{P}] + [\mathrm{C}] + K_\mathrm{D})^2 - 4\,[\mathrm{P}]\,[\mathrm{C}]}}{2}. \tag{5.1}$$

Under assumption of a K_D of 200 nM[17], the curves in Figure 5.10 were obtained for the Sec7 domain of ARNO and SecinH3.

For the planned experiments it was important to maximise the portion of labelled protein. So, with the intent to minimize unspecific interaction, minimal concentrations

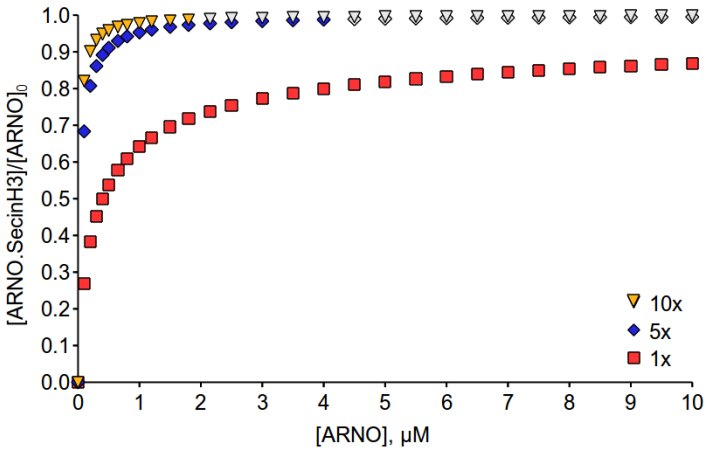

Figure 5.10: Simulation of the theoretical binding ratio for a K_D of 200 nM - The ratio of SecinH3:ARNO-Sec7 complex with respect to total ARNO-Sec7 concentration is plotted against ARNO-Sec7 concentration. The results for 1-fold (red square), 5-fold (blue diamond) and 10-fold (yellow triangle) compound excess are shown. Data points which require compound concentrations above SecinH3 solubility are shown in grey. With 1 µM ARNO-Sec7, 5-fold excess SecinH3 is expected to lead to 95 % complexed protein at equilibrium.

giving acceptable complex ratio were chosen. The preliminary labelling experiments were therefore performed at [P] = 0.5 µM and [C] = 2.5 µM.photo-activatable group

5.3.2.2 Correction of the calculations for a K_D of 15 µM

Later on, new binding data were obtained which estimated the dissociation constant of ARNO-Sec7 and SecinH3 at ∼15 µM (A. Schmitz, unpublished data). The corrected curves in Figure 5.11 show that with the concentrations used above a maximal labelling of only 20 % is possible. The easiest way to improve this ratio, would be increasing the compound concentration. Unfortunately, SecinH3 was known to be only limitedly soluble, though the exact solubility had not been determined. Thus, the solubility of SecinH3 was measured.

5.3.2.3 Solubility of SecinH3

The solubility of small molecules was determined by measuring the absorption of compound dilutions after centrifugation to remove possible aggregates/precipitates. The solubility was defined as the concentration at which the linear correlation between

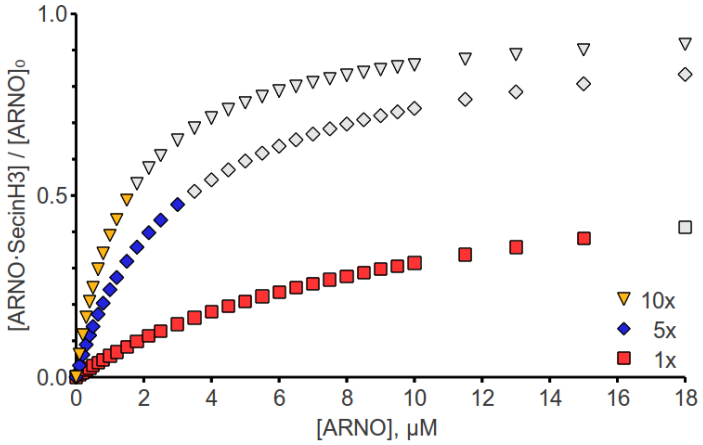

Figure 5.11: Simulation of the theoretical binding ratio for a K_D of 15 µM - The ratio of SecinH3:ARNO-Sec7 complex with respect to total ARNO-Sec7 concentration is plotted against ARNO-Sec7 concentration. The results for 1-fold (red square), 5-fold (blue diamond) and 10-fold (yellow triangle) compound excess are shown. Data points which require compound concentrations above SecinH3 solubility are shown in grey.

concentration and absorption is broken. Since SecinH3-TPD is not stable at the wavelengths used, the solubility tests were performed with SecinH3 instead.

Traditionally, the SecinH3 inhibition experiments *in vitro* and in cell culture were performed in 0.4 – 10 % DMSO in aqueous solution. Though higher DMSO concentration could improve SecinH3 solubility, it is known that most proteins do not withstand it. Hence, as an alternative to DMSO, diglyme (1-methoxy-2-(2-methoxyethoxy)ethane) was tested. Diglyme is an organic solvent similar to glycerol, which has been shown to be tolerated by various enzymes even at high concentrations[124–126].

The absorption spectrum of SecinH3 in diglyme was measured and an absorption maximum was found at 260 nM (Fig. 5.12). Since the absorption maximum in DMSO was known to be at 280 nM, the measurements were performed at this wavelength. Because of the high absorption at 280 nM, the solubility curves at high concentrations were measured at 330 nM.

As visible in Figure 5.13**a**, SecinH3 has a solubility of ∼20 µM both in 10 % DMSO and diglyme. Desirably, the solubility in 40 % diglyme drastically increases to over 100 µM. Even at 20 % diglyme, SecinH3 is soluble till 100 µM (Fig. 5.13**b**).

Because no data about the tolerance of ARNO for diglyme were available, both diglyme and DMSO were initially used for the labelling experiments.

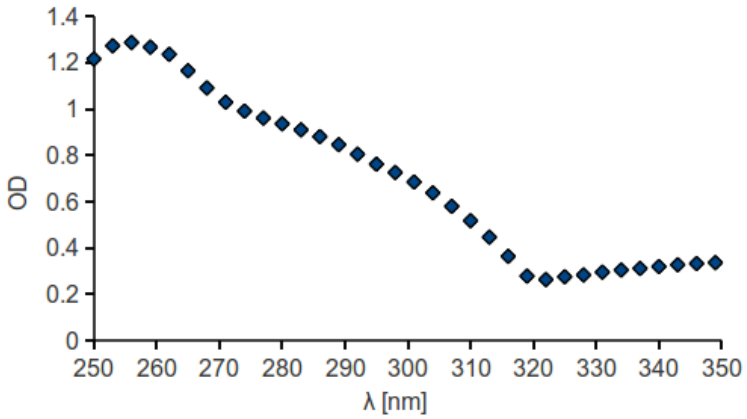

Figure 5.12: Absorption spectrum of SecinH3 in diglyme - Absorption of a 50 µM solution of SecinH3 in diglyme was measured

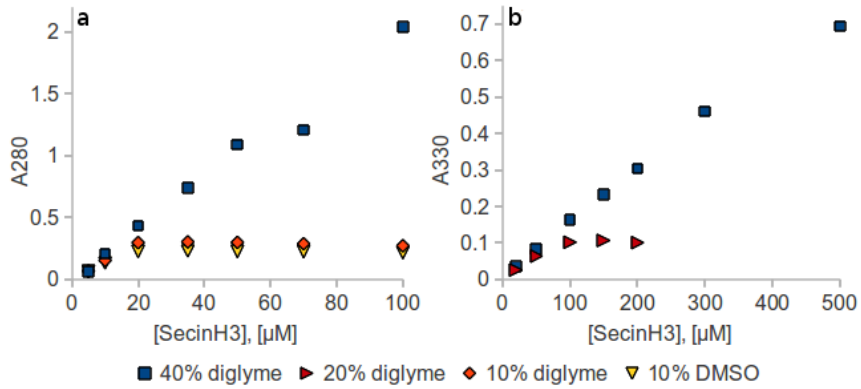

Figure 5.13: Diglyme increases SecinH3 solubility - SecinH3 dilutions in 10 % (orange diamond), 20 % (red triangle) or 40 % (blue square) diglyme or 10 % DMSO (yellow triangle) were centrifuged to remove non soluble compound. Absorption at 280 (**a**) or 330 (**b**) nm is plotted against SecinH3 concentration. The break-point in the curves indicate the maximal solubility. The solubility of 20 µM in 10 % DMSO is increased to 100 µM in 20 % diglyme.

5.3.2.4 Analysis of the specificity of labelling

The most important question at this stage was whether the synthesised SecinH3 probe would label ARNO-Sec7 and how specific this labelling would be. As SecinH3-TPD carries a desthiobiotin tag, the proteins it reacts with can easily be detected on a western blot by streptavidine. As a first test of specificity, labelling of the purified PAZ domain of Argonaute1 was measured. This protein was chosen as negative control from those available in the group, because it has similar size and carries the same tag (His)[73,74] as ARNO-Sec7, allowing easy comparison of the total protein in the assay.

Figure 5.14 shows that SecinH3-TPD labelled ARNO-Sec7 upon irradiation. No signal was detected when the negative control compound BA103 (Fig. 5.5) was used. Although the PAZ domain was slightly tagged by SecinH3-TPD, the binding was almost completely competed by ARNO-Sec7. These results show that the designed SecinH3 probe retains its affinity for ARNO.

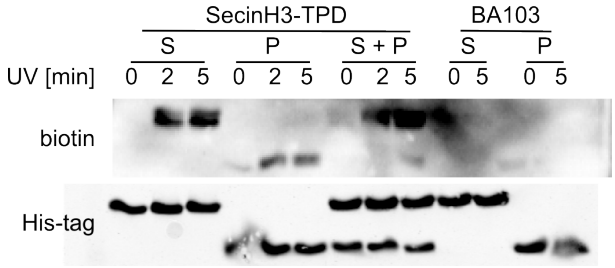

Figure 5.14: SecinH3-TPD binds and labels ARNO-Sec7 - SecinH3-TPD or the negative control compound BA103 (2.5 µM) were incubated with the indicated proteins (0.5 µM) for 15 min at RT before irradiation for 0, 2, or 5 min. After separation with 10 % SDS-PAGE and Western blotting, the labelled proteins were detected with a horseradish peroxidase streptavidin conjugate. The total amount of protein was detected with His_5 specific antibody. S: ARNO-Sec7, P: PAZ-domain

Having seen that SecinH3-TPD binds to ARNO-Sec7, I tested if the labelling efficiency could be improved with higher compound concentrations. To allow the use of concentrations higher than 20 µM, diglyme was used as solvent. As reference, 20 µl SecinH3-TPD in DMSO was used. Diglyme increased dramatically the amount of labelled ARNO-Sec7. As indicated by the bands intensities in Figure 5.15, the amount of modification was improved even at equal compound concentration as in DMSO (20 µM). Increasing the concentration to 80 µM had a less prominent effect.

Two additional main differences are evident in the diglyme samples. First, the appearance of additional bands: the presence of bands at around the double mass of ARNO-Sec7 was observed also in other experiments and seems to be quite condition dependent. Second, a band shift in the irradiated samples (His-tag detection): the expected mass difference in case of single labelling (∼700 Da) is not a sufficient explanation, in particular because the band shift is not present in the DMSO samples. A possible interpretation, which would also account for the higher labelling observed at 20 µM, is that diglyme lead to multiple labelling of a single ARNO molecule. Taken together with the fact, that diglyme hinders the nucleotide exchange of ARNO-Sec7 on ARF1 in vitro (A. Bill, unpublished data), diglyme does not seem to be compatible with the ARNO-Sec7 activity.

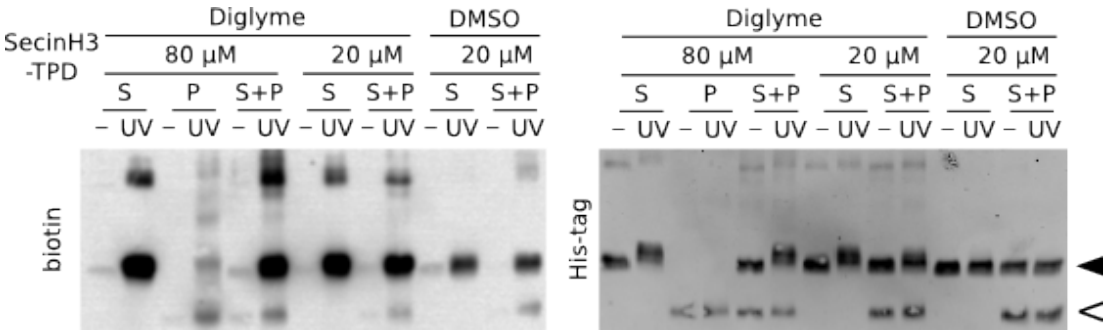

Figure 5.15: Diglyme increases ARNO-Sec7 labelling by SecinH3-TPD - The labelling experiments were performed in 20 % diglyme or 10 % DMSO as indicated. The proteins (1 µM each) were incubated for 15 min at RT with the given concentrations of inhibitor and then either irradiated for 10 min (UV) or kept in the dark (-). After separation on 12.5 % SDS-PAGE, detection was performed as in Figure 5.14. S: ARNO-Sec7 (◀), P: PAZ-domain (<)

5.3.2.5 Labelling efficiency of ARNO domains and mutants

At this point, only labelling of a truncated form of ARNO, namely ARNO-Sec7, had been investigated. Yet, although SecinH3 was selected against the Sec7 domain, its biological activity imply that SecinH3 is successfully inhibiting the native protein. Therefore I was interested in finding out how the different ARNO domains influence SecinH3 binding. Additionally to the full length ARNO (FL), constructs lacking the pleckstrin homology and coiled coil domains (ΔPH and ΔCC, respectively) were used. Moreover, I made a new construct lacking the polybasic region at the C-terminus (ΔPBR). As

described in Section 1.2.1.1, p. 6, this region is responsible for autoinhibition of the exchange activity of the Sec7 domain and its excision restores the activity in purified proteins[19]. Figure 5.16 shows that SecinH3-TPD labelled ΔPBR much better than ARNO-Sec7. Although the actual concentration of the different domains in the assay was quite variable, ARNO-Sec7 and ΔPBR were present in comparable concentrations.

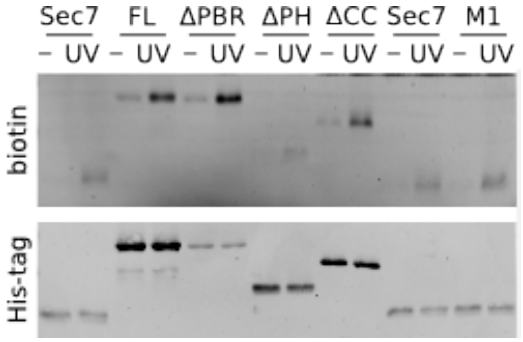

Figure 5.16: ΔPBR is more efficiently labelled than ARNO-Sec7 - The proteins (1 µM) were incubated with 20 µM SecinH3-TPD for 10 min on ice before irradiation (10 min). After separation on a 12.5 % SDS-PAGE and Western blotting, the labelled proteins were detected with a NeutrAvidin DyLight 800 fluorescent conjugate. The total amount of protein was detected with His_5 specific antibody. The abbreviations are defined in the text. -: non irradiated; UV: irradiated for 10 min.

Molecular docking experiments performed in the group of Prof. Bajorath (Department of life science informatics, University of Bonn), revealed a possible binding site of SecinH3 on ARNO-Sec7. On this basis, four point mutants which should affect SecinH3 binding were designed (Fig. 5.17). Though the mutants were functional in the exchange assay, the activity was too low to permit the measurement of a significant inhibition assay (Caroline Kubaczka, unpublished data). Thus, the inhibition (ergo binding) of SecinH3 of this mutants could not be tested that way. A more direct approach to analyse their binding to SecinH3 was the use of the photoreactive SecinH3. Surprisingly, none of the mutants showed a reduced binding. Yet, for the mutant L148A (M1) the labelling was even increased (Fig. 5.16).

Detection with His_5 antibody revealed that the actual concentration of the different proteins in the assay was not comparable. However, the relative concentrations in the repetitions of the assay were not constant. This excludes an error in the estimation protein concentration. To check if aggregation or coating to the vial were the causes

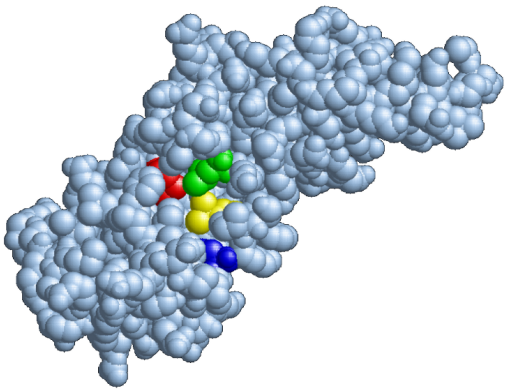

Figure 5.17: Mutants 1-4 of ARNO-Sec7 - A possible binding site of SecinH3 on ARNO-Sec7 was detected with the help of docking experiments *in silico*. The four point mutants which were designed to disrupt the interaction are shown in color in the model of the ARNO-Sec7 domain. Yellow: M1, L148A; green: M2, L153A; red: M3, M164A; blue: M4, F243A. Image based on Protein Data Bank entry 1PBV (Ref. 11).

of variability, the samples were centrifuged before loading on the gel. Both the supernatant, containing the soluble proteins, and the pellet, containing aggregated proteins, were analysed. Figure 5.18**a** shows that most of the protein was actually found in the pellet, demonstrating a solubility issue.

5.3.2.6 Effect of detergent on solubility and labelling

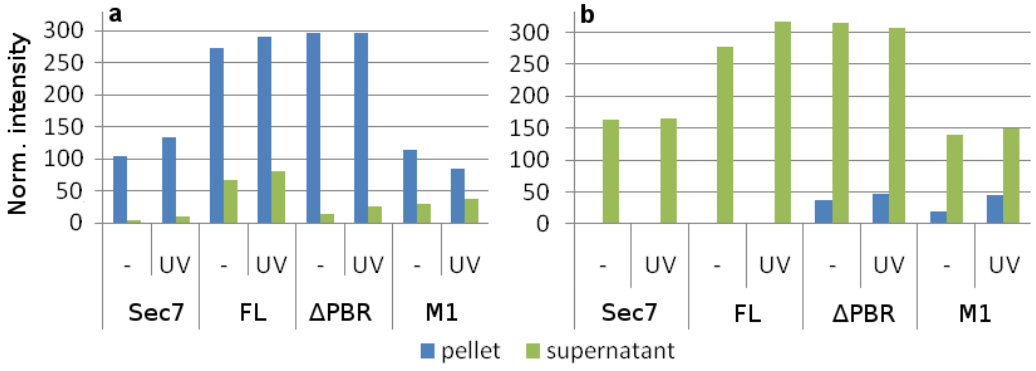

Figure 5.18: Triton X-100 is necessary to keep ARNO in solution - Quantification of total protein is displayed. The experiments were performed as in Figure 5.16, with the difference that after irradiation the samples were centrifuged to separate soluble (supernatant, green) from aggregated (pellet, blue) proteins. **a**. Labelling under standard conditions **b**. 0.005 % Triton X-100 added to the labelling buffer.

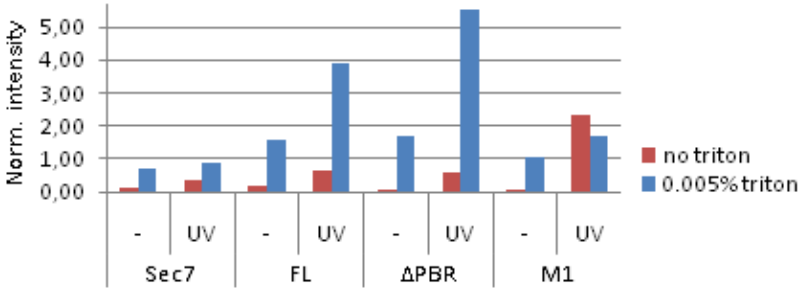

Figure 5.19: The amount of labelled ARNO in the supernatant is increased by Triton X-100 - Quantification of labelled protein found in supernatant in the experiment in Figure 5.18.

With the intent to improve solubility, 0.005 % Triton X-100, a nonionic surfactant which is often used to solubilise proteins, was added to the labelling buffer. Indeed, the detergent improved the solubility dramatically and most of the protein was found in solution (Fig. 5.18**b**). As expected, the amount of labelled protein in the supernatant was also increased (Fig. 5.19). Under this conditions, the labelling of ΔPBR was 5-times higher than ARNO-Sec7 labelling. The higher background signal for non irradiated ΔPBR is due to non specific recognition of ΔPBR by NeutrAvidin.

5.3.3 Establishment of the mass spectrometric analysis

Having proved that SecinH3-TPD can be crosslinked to ARNO, the next step was to define suitable conditions for the mass spectrometric analysis of ARNO. Since there was no know-how about protein analysis on MS in the group, we decided to start with a model sample before moving to more complex mixtures.

The whole procedure from the labelled protein, through LC-MS/MS analysis, till peptide identification involves various steps like denaturation, digestion, HPLC (high pressure liquid chromatography) and MS. Hence, we decided to synthesise a peptide which would allow to train all these techniques.

The ideal peptide, would be similar to those expected for digested ARNO and carry at least one cleavage site for trypsin. Thus, the following 27 amino acids long sequence was chosen out of ARNO-Sec7 sequence ("‖" denotes a tryptic cleavage site):

$$H_2N - FVAMNR \parallel GINEGGDLPEELLR \parallel NLYDSIR - CO_2H$$

5.3.3.1 Synthesis of a reference peptide

The synthesis was carried out on a peptide synthesiser by Fmoc solid-phase peptide synthesis. Deprotection was achieved in 82.5 % TFA, 2.5 % DTT, 5 % water, 5 % phenol and 5 % thioanisol as described in Section 7.1.1.2 (p. 123). Following HPLC purification, the identity of the product was confirmed by LC-MS and MS (Fig. 5.20).

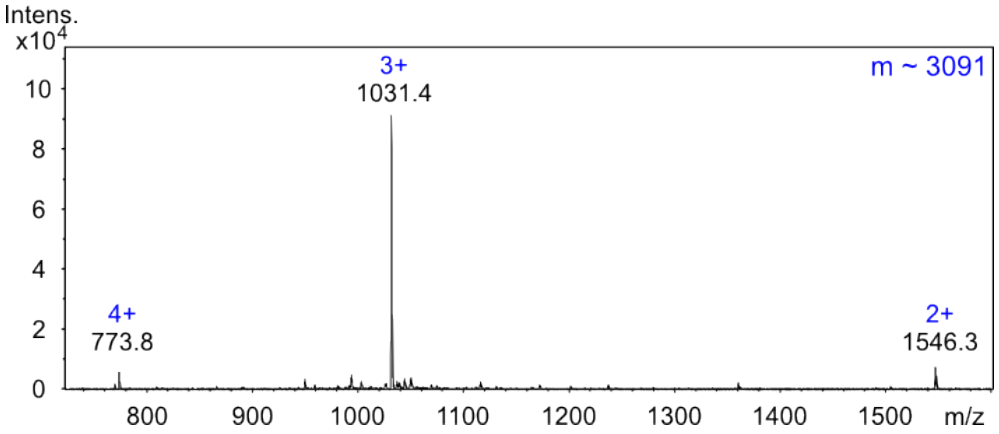

Figure 5.20: Mass spectrum of the synthesised reference peptide - The identity of the product was confirmed by MS. Expected mass: 3091.4 g/mol.

5.3.3.2 Optimisation of the tryptic digestion

Proteins need to be denaturated before digestion to expose the cleavage sites to the proteases. Even if the reference peptide was not expected to assume secondary structures which could disturb cleavage, it was important to test the whole process, to make sure that the denaturation conditions would non inhibit the subsequent digestion. Thus, various conditions were tested for denaturation and digestion of the peptide.

Denaturation in 0.1 % SDS with subsequent digestion with 1:100 w/w trypsin gave good results (Fig. 5.21) but has the disadvantage of SDS non being compatible with reverse phase (RP) chromatography. Even though purification with ZipTip before HPLC worked for the peptide, their use with ARNO-Sec7 was not satisfying.

Boiling in urea, was found to lead to modification of the peptide N-terminus. As shown in Table 5.1, a difference of 42-43 Da between expected and measured mass was detected both in the undigested reference peptide and in the N-terminal peptide. This difference can be explained by the reaction of the free amine with urea or its decompositions products, as ammonium cyanate (Fig. 5.22**a** and **b**, respectively). Indeed,

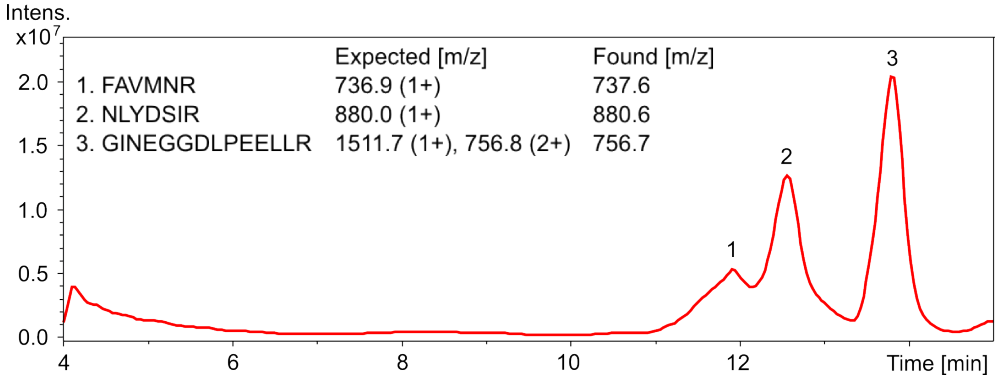

Figure 5.21: Total ion chromatogram of the trypsinated reference peptide - The reference peptide was digested with trypsin and analysed by LC-MS. All the three expected fragment could be identified (insert).

Table 5.1: Boiling in urea causes N-terminus modification - A difference of 42-43 Da was found for the N-terminal peptide.

		m/z		
Peptide	**z**	**Expected**	**Found**	**Δm [Da]**
Undigested	3	1031.8	1045.7	41.6
	4	774.1	784.6	41.9
FVAMNR	1	737.4	780.4	43.0

the formation of cyanic acid in aqueous urea at high temperatures, and the subsequent carbamylation of the N-terminus and lysine, arginine, and cysteine residues was reported[127–130]. To prevent the side reactions, which can heavily disturb protein identification, one should reduce the denaturation temperature (and time) and use fresh prepared urea solution, so as to reduce the presence of cyanic acid to a minimum.

As a third method, which does not require SDS nor urea, acetone precipitation was tested. Addition of an excess of acetone denaturates and precipitates the proteins in the sample. Digestion is then performed after (partial) reconstitution in ammonium bicarbonate[131]. This approach was directly tested on ARNO-Sec7 and worked well enough to allow to start with optimisation of the LC-MS conditions.

Particularly interesting for the analysis of SecinH3-TPD labelled samples, was the digestion on filter[129]. This protocol, where the whole sample preparation process is performed on common ultrafiltration devices, circumvents any solubility issue and offers at the same time a nice method to get rid of excess compound. The proteins are

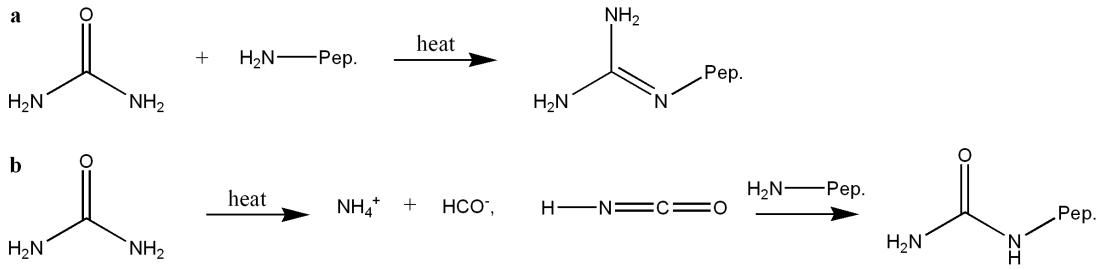

Figure 5.22: Possible reactions of urea with the N-terminal amine - The reaction in **a** leads to an increase in mass of 42 Da. Reaction **b** adds 43 Da to the peptide total mass.

denatured and solubilised in 0.1 % SDS, which is then exchanged by urea on a standard filtration device. At this stage, also unbound compound is washed away. Incubation of a trypsin solution on the filters leads to digestion of the proteins and the resulting peptides are then collected by centrifugation.

5.3.3.3 Optimisation of the LC-MS conditions

LC-MS analysis of the reference peptide digest was performed on a reversed-phase chromatography C4 column and could be optimised as desired (Fig. 5.21). For better separation of the more complex mixture obtained by digestion of ARNO-Sec7, the gradient was extended to 60 min. Reduction of the flow rate from 0.4 ml/min to 0.2 ml/min improved the separation further. The C4 column already allowed good separation of the peptides, but since some were coeluting, a C18 column was then used, with which good chromatograms could be acquired (Fig. 5.23). At the end of the optimisation process, I was able to identify 17 of the 21 expected completely digested tryptic peptides, with a sequence coverage of 92 % (Fig. 5.24).

5.3.3.4 Tandem LC-MS measurements

The LC-MS measurements were analysed manually and are only suited for samples with limited complexity, such as digests of purified proteins. For the analysis of more complex samples from multiple and unknown proteins, I tested the conditions for tandem LC-MS. As tandem MS delivers sequence information additionally to the peptide mass, it is more powerful in the automated analysis of unknown samples. After optimisation of

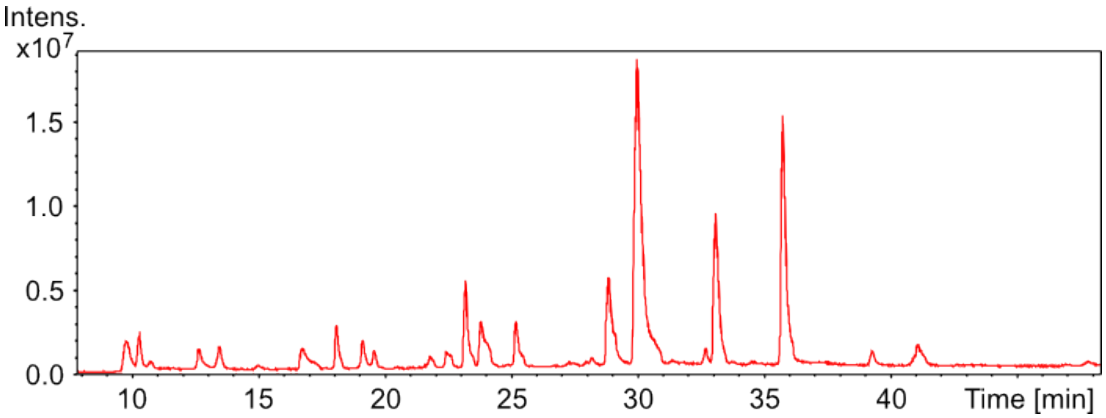

Figure 5.23: Total ion chromatogram of an ARNO-Sec7 digest - ARNO-Sec7 was digested with trypsin after acetone precipitation. Separation on a RP C18 column with a 60 min gradient lead to good separated peaks.

1 GSSHHHHHHS SGLVPRGSHM EANEGSKTLQ RNRKMAMGRK KFNMDPKKGI

51 QFLVENELLQ NTPEEIARFL YKGEGLNKTA IGDYLGEREE LNLAVLHAFV

101 DLHEFTDLNL VQALRQFLWS FRLPGEAQKI DRMMEAFAQR YCLCNPGVFQ

151 STDTCYVLSF AVIMLNTSLH NPNVRDKPGL ERFVAMNRGI NEGGDLPEEL

201 LRNLYDSIRN EPFKIPEDDG NDVTHS

Figure 5.24: ARNO-Sec7 was analysed with good sequence coverage - Sequence of ARNO-Sec7 with the peptides found in the LC-MS run (Fig. 5.23) highlighted. Black: completely digested peptides; blue: peptides with one missed cleavage.

the separation gradient and the MS/MS parameters, good peptides scores and sequence coverage were obtained for both BSA and ARNO-Sec7 samples.

5.3.4 Detection of modified proteins and fragments by LC-MS

To test if the analytical conditions enable in principle the identification of protein modification, I alkylated some ARNO-Sec7 with iodoacetamide (IAA).

5.3.4.1 Analysis of iodoacetamide modified ARNO-Sec7

IAA reacts with the side chain of cysteine leading to a mass increase of 57 Da (Fig. 5.25). Since ARNO-Sec7 sequence has three cysteines, a mass difference of 171 Da is expected for the fully alkylated protein.

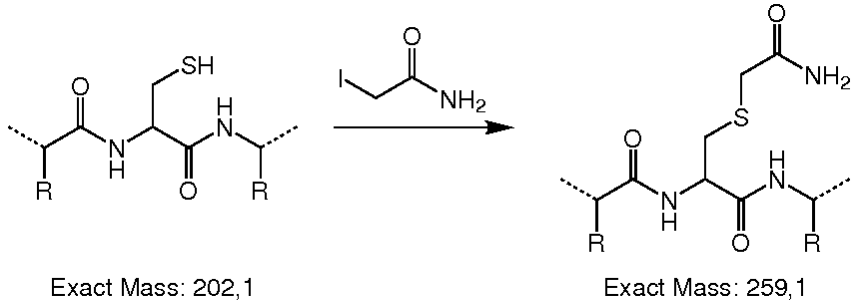

Figure 5.25: Reaction of iodoacetamide with cysteine side chains - A mass increase of 57 Da is observed upon alkylation with IAA.

Indeed, LC-MS analysis of non digested ARNO-Sec7, revealed a mass increase of ∼182 Da in the sample treated with IAA (Fig. 5.26) and confirmed that the method was suitable for the identification of small modifications. The small deviation between the theoretical and experimental mass lies in the resolution of the instrument.

The same samples were then analysed after digestion. All cysteines in ARNO-Sec7 are found in the same tryptic peptide, for which a mass difference of 171 Da is again expected. As visible in Figure 5.27, some additional peaks appeared in the chromatogram after alkylation. Some of them were peptides which occurred in higher amount without carrying any alkylation, perhaps due to improved digestion of the alkylated sample. More interestingly, the peak at 44 min could be attributed to the three-times alkylated peptide (Fig. 5.28**b**). As a proof of the sensitivity of the system,

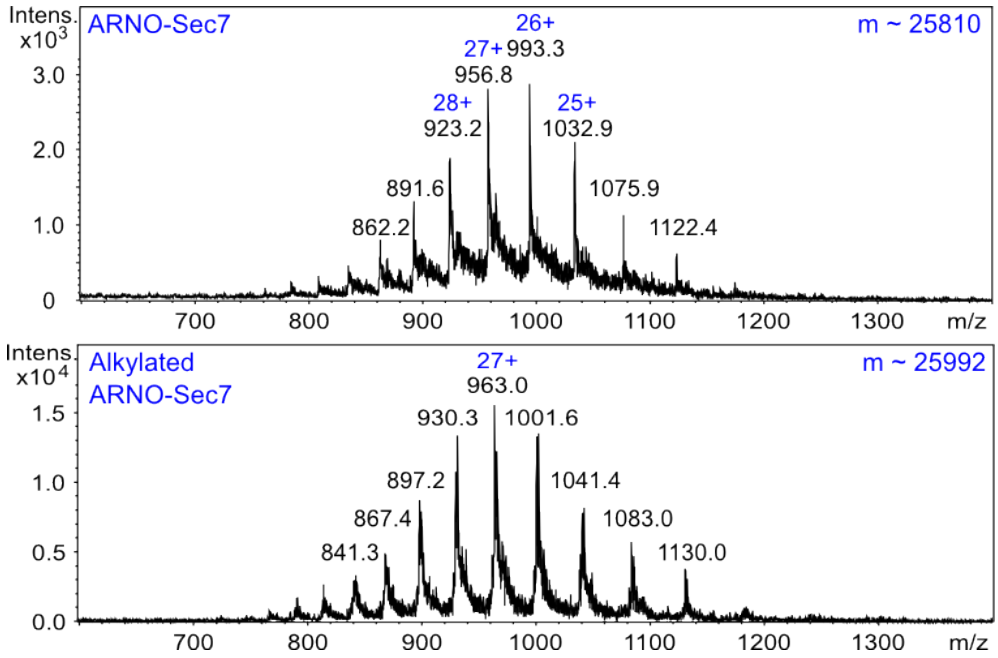

Figure 5.26: The mass increase upon alkylation can be identified in the non digested ARNO-Sec7 - ARNO-Sec7 was alkylated with IAA and analysed by LC-MS (lower panel). The average mass spectrum at the protein elution time is shown. The mass spectrum of unmodified ARNO-Sec7 is shown as reference (upper panel). The mass difference observed is in the range expected for a three-time alkylation.

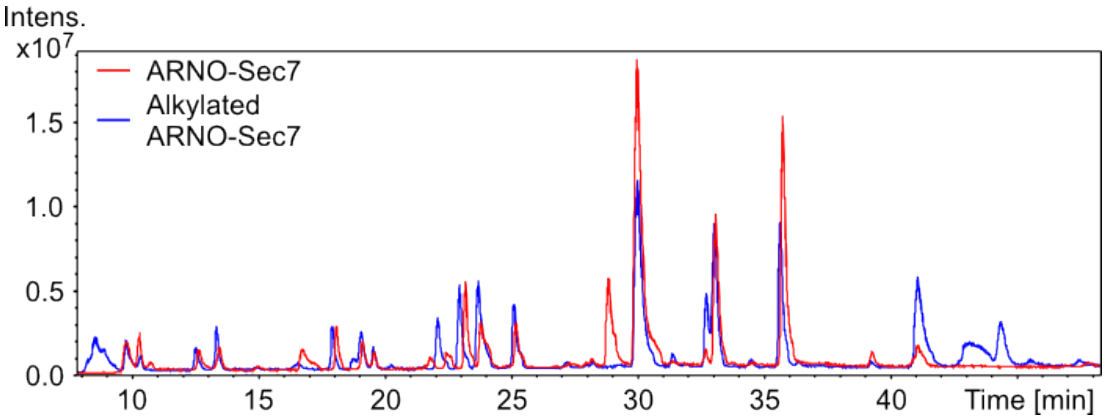

Figure 5.27: Total ion chromatogramm of alkylated ARNO-Sec7 - Comparison of LC-MS analyses of alkylated and unmodified digested ARNO-Sec7. Some additional peaks were observed in the alkylated sample (blue).

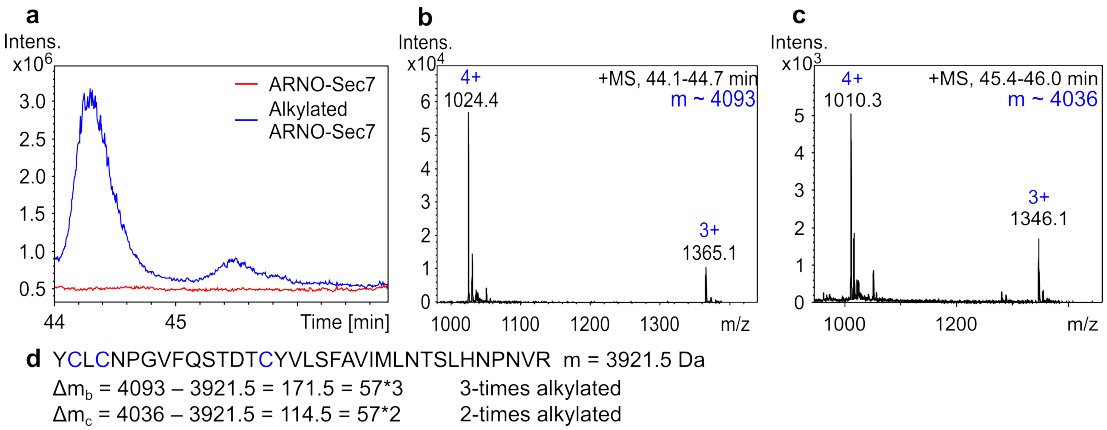

Figure 5.28: Additional peaks correspond to three- and two-times alkylated peptides - a. Magnification of the 44-47 min window of the chromatogram in Figure 5.27. **b, c.** The mass spectra relative to the two peaks identify a three- (**b**) and two-times (**c**) alkylated peptide. **d.** Sequence of the alkylated peptide and mass differences measured.

a low intensity peak at 45 min, which corresponded to a small amount of two-times alkylated peptide, was also identified (Fig. 5.28**c**).

5.3.4.2 Analysis of SecinH3-TPD-labelled ARNO-Sec7

Since detection of protein alkylation was successful, I tested if it was possible to detect labelling by SecinH3-TPD. ARNO-Sec7 was incubated either with SecinH3-TPD, the negative control compound BA103, or DMSO before irradiation with UV. Sample preparation was performed both with the acetone precipitation and the digestion on filter methods. Unfortunately, no relevant difference was seen between the SecinH3-TPD samples and the negative controls (data not shown). It is possible that SecinH3 was bound to a fragment which is not detected with the analysis conditions or, more probably, that the portion of labelled peptide was to low to be detected.

5.3.5 Enrichment of biotinylated ARNO

A strategy to improve detection of the labelled protein, is to selectively enrich it. Since SecinH3-TPD carries a biotin, covalently bound protein can be fished with streptavidine conjugated beads. This method allows enriching biotinylated protein from very complex sample, and is therefore ideal to test labelling specificity in cell lysate, too (see Section 5.1, p. 65, for explanation of the CCMS strategy). The enrichment protocol

was adapted from Weerapana *et al.*[92] and first tested with prebiotinylated ARNO-Sec7 (bio-Sec7).

Table 5.2: Enrichment and identification of bio-Sec7 - bio-Sec7 (250 nM), pure or mixed to 2 mg/ml H460 proteome, was enriched with streptavidine beads, analysed by LC-MS/MS and the data searched against the SwissProt database using Mascot. Samples containing non biotinylated ARNO-Sec7 served as control for unspecific binding to the beads. ARNO was only identified in samples containing bio-Sec7. No other proteins were identified. The proteome did not disturb identification.

Sample	Proteins identified	Score
ARNO-Sec7	none	-
bio-Sec7	ARNO	90
Proteome + ARNO-Sec7	none	-
Proteome + bio-Sec7	ARNO	100

5.3.5.1 Determination of the detection limit

Streptavidine beads were incubated with samples containing either bio-Sec7 or ARNO-Sec7 only, or in combination with cell lysate. After intensive washing, the enriched proteins were digested on the beads. The supernatant was analysed by LC-MS/MS and the data searched against the SwissProt database using Mascot[132].

Table 5.3: Determination of the detection limit for bio-Sec7 - The experiment in Table 5.2 was repeated with decreasing concentrations of bio-ΔPBR. Identification of ARNO was possible even with the lowest starting concentration (50 nM). Again, the proteome did not disturb identification.

Sample	Proteins identified	Score
250 nM bio-ΔPBR	ARNO	459
100 nM bio-ΔPBR	ARNO	192
50 nM bio-ΔPBR	ARNO	136
Proteome + 1 µM ΔPBR	none	-
Proteome + 100 nM bio-ΔPBR	ARNO	252

Indeed, ARNO was only identified in the samples containing bio-Sec7 and the presence of lysate did not disturb identification (Tab. 5.2). Unfortunately, it was not possible to detect the biotinylated peptide after elution from the beads (data not shown).

The first enrichment experiment was performed with relatively high concentration of bio-Sec7 (250 nM). To check the detection limit of the method with respect to the start concentration of labelled protein, the procedure was repeated with a dilution series. For this experiment, prebiotinylated ΔPBR (bio-ΔPBR) between 250 and 50 nM to was used. Even in the 50 nM bio-ΔPBR sample (which corresponds to around 5 % labelling of 1 µM ARNO), it was still possible to identify ARNO and again, the presence of lysate did not influence the detection in the 100 nM sample (Tab. 5.3). This shows that the protocol is suitable to selectively enrich biotinylated proteins even in relatively low concentration and in complex samples.

5.3.6 Specificity of labelling in cell lysate

Having a method to analyse tagged proteins in cell lysate, I moved to the determination of specificity of SecinH3-TPD labelling in cell lysate. For this, it was necessary to find lysis conditions which do not disturb the labelling reaction.

5.3.6.1 Determination of lysis conditions compatible with labelling

Triton X-100 and digitonin, two detergents which are often used for cell lysis, were tested in the labelling assay. As evident in Figure 5.29, Triton completely blocks labelling, while digitonin does not affect the reaction. Thus, I decided to prepare the proteome by permeabilising the cells with 25 µg/ml digitonin[133].

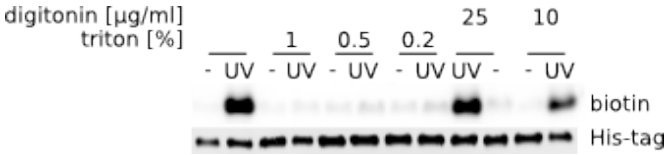

Figure 5.29: Digitonin is compatible with SecinH3-TPD labelling - Triton X-100 and digitonin were added to the labelling buffer in the indicated concentrations and the labelling assay performed as usual. Triton inhibits while digitonin does not affect labelling.

5.3.6.2 Analysis of specificity by Western blot

To test if ARNO-Sec7 is labelled by SecinH3 also in cell lysate, HEK (Human Embryonic Kidney 293) cells were harvested, permeabilised and the soluble proteins collected by centrifugation. The labelling reaction was performed in diluted proteome (1.1 mg/ml total protein) with or without addition of purified ARNO-Sec7 (1 µM) and with 20 µM

SecinH3-TPD in 20 % diglyme. Promisingly, a band was detected at the right size (Fig. 5.30**a**, lane 6). Few additional bands were detected at higher size. In particular, a strong signal at ∼70 kDa was present only in the sample where no ARNO-Sec7 was added (lane 8).

As a second experiment, I looked if labelling of ARNO could be competed by the addition of unmodified SecinH3. H460 cells were prepared as above and the labelling reaction was done in proteome diluted to 2 mg/ml total protein. 1 µM ΔPBR, 5 µM SecinH3-TPD and 15 µM SecinH3 in 10 % DMSO were used. Again, labelling of ΔPBR was possible also in cell lysate (Fig. 5.30**b**, lane 14). Interestingly, like in the precedent experiment with no ARNO addition, competition with SecinH3 reduced drastically the amount of labelled ΔPBR and a new strong band at ∼80 kDa was observed (lane 16). Surprisingly, this band was not present in the sample without ΔPBR (lane 12).

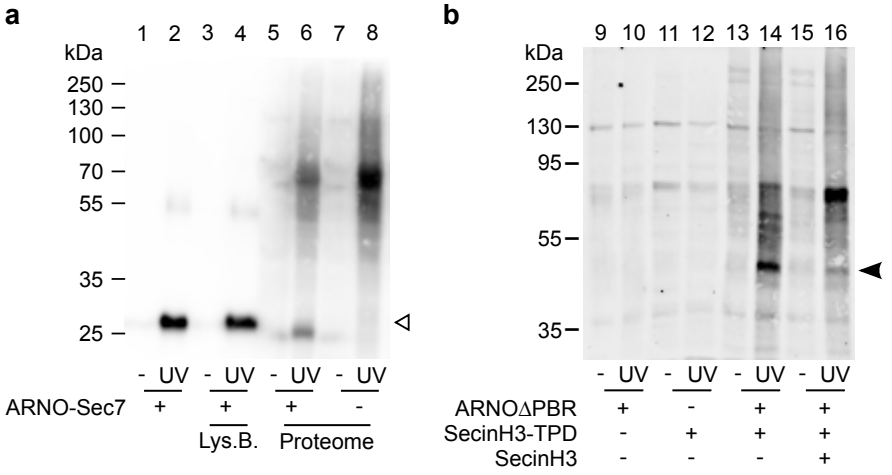

Figure 5.30: ARNO can be labelled with SecinH3-TPD in cell lysate - Both ARNO-Sec7 and ΔPBR were labelled by SecinH3-TPD in cell lysate (lanes 6 and 14, respectively). Interestingly, when no ARNO-Sec7 was added (lane 8) or binding to ΔPBR was competed with unmodified SecinH3 (lane 16), a band at around 70-80 kDa appeared instead of the ARNO band. **a.** The standard labelling reaction (lane 2) was compared with reactions were lysis buffer (Lys.B., lane 4) or cell lysate (Proteome, lane 6) were added. Labelling of proteome without addition of purified ARNO-Sec7 (1 µM) is visible in lane 8. After separation on a 10 % SDS-PAGE and Western blotting, the labelled proteins were visualised with horseradish peroxidase streptavidin conjugate. 20 µM SecinH3-TPD in 20 % diglyme were used. ◁: ARNO-Sec7. **b.** Labelling by SecinH3-TPD in cell lysate, was competed with underivatised SecinH3. After separation on a 7.5 % SDS-PAGE and Western blotting, the labelled proteins were detected with NeutrAvidin DyLight 800 fluorescent conjugate. 1 µM ΔPBR, 5 µM SecinH3-TPD and 15 µM SecinH3 in 10 % DMSO were used. ◀: ΔPBR.

Table 5.4: Human proteins binding biotin - Database search revealed three biotin-binding proteins with size similar to that of the unknown protein labelled by SecinH3-TPD.

UniProt accession number	Protein	Size (kDa)
P50747	Biotin–protein ligase	80.8
Q96RQ3	Methylcrotonoyl-CoA carboxylase subunit alpha, mitochondrial	80.5
P05165	Propionyl-CoA carboxylase alpha chain, mitochondrial	80.1

When proteome analysis with photoreactive probes are performed, competition is used to select which bands to analyse[87,134–136]. Indeed, a main disadvantage of photoreactive affinity based probes over activity based probes is the higher background labelling which mostly leads to detection of multiple bands even for specific inhibitors. The idea behind the competition experiments is that only specific labelling will be drastically reduced by the unmodified inhibitor. Thus, often only the bands with reduced labelling are analysed further[87,134–136].

The experiment in Figure 5.30**b** clearly shows that labelling of ΔPBR is indeed due to interaction with SecinH3. However, some other bands are present which are competed by SecinH3. The MS analysis of all labelled proteins is described in next Section.

The ∼70-80 kDa band is an unexpected result. Indeed, I am not aware of competition experiments with photoreactive probes which lead to such a drastic increase of signal. Since the labelling is higher in the presence of unmodified SecinH3, it can not be attributed to a specific interaction with the inhibitor. However, binding is strong enough to lead to very high signal. In fact, the interaction between SecinH3-TPD and the unknown protein, could also be mediated by the photoreactive moiety or the linked desthiobiotin. In this case, SecinH3 would not disturb the binding.

Since biotin is naturally present in cells, SecinH3-TPD could possibly interact via its reporter group with some biotin binding protein. Searching the UniProtKB database for human proteins with a biotinyl-binding domain gave 5 hits, two of which with size in the right range. The QuickGO* browser delivered an additional ∼80 kDa human

*QuickGO is a fast web-based browser for Gene Ontology terms and annotations, which is provided by the UniProtKB-GOA group at the European Bioinformatics Institute (EBI). Adress: http://www.ebi.ac.uk/QuickGO/

Table 5.5: Identification of ARNO by CCMS - ΔPBR was labelled with SecinH3-TPD, enriched, and analysed by LC-MS/MS. Data base search with Mascot identified one relevant peptide. When the same experiment was performed in the presence of proteome, ARNO was not identified.

	Experimental			Calculated			
Name	**m/z**	**z**	**m**	**m**	**Δm**	**Score**	**Sequence**
CYH2_HUMAN Cytohesin-2 Mass: 46859 **Score: 66**	756.5	2	1511.0	1510.8	0.2	**53**	GINEGGDLPEELLR

protein with biotin binding function (search ID: GO:0009374). These proteins are shown in Table 5.4. It is therefore possible that SecinH3-TPD, when displaced from ARNO by SecinH3, or when only very low amounts of ARNO are present, binds and labels biotin-binding proteins through its reporter group. However, this hypothesis has not been verified.

5.3.6.3 Identification of interacting proteins by MS

To identify the proteins labelled by SecinH3-TPD, the complete CCMS procedure was followed (see Section 5.3.5, p. 91). The labelling reaction was performed with or without proteome from H460 cells, with the addition of 1 μM ΔPBR.

Table 5.6: Identification of tubulin by CCMS - H460 proteome was labelled with SecinH3-TPD, enriched, and analysed by LC-MS/MS. Data base search with Mascot identified various tubulin isotypes with good scores and sequence coverage. The identified proteins and peptides are listed with the relative scores.

	Experimental			Calculated			
Name	**m/z**	**z**	**m**	**m**	**Δm**	**Score**	**Sequence**
TBB5_HUMAN	566.0	2	1130.0	1129.6	0.4	**32**	FPGQLNADLR
Tubulin	572.5	2	1143.0	1142.6	0.4	**55**	LAVNMVPFPR
beta chain	615.5	2	1229.0	1228.6	0.4	**87**	ISEQFTAMFR
Mass: 50095	651.5	2	1301.0	1300.6	0.4	**53**	ISVYYNEATGGK
Score: 431	660.5	2	1319.0	1318.7	0.3	**61**	IMNTFSVVPSPK
	723.9	2	1445.8	1445.7	0.1	**58**	EVDEQMLNVQNK
	808.6	2	1615.2	1614.8	0.4	**40**	AILVDLEPGTMDSVR
	830.7	2	1659.4	1658.9	0.5	**35**	ALTVPELTQQVFDAK
	608.5	3	1822.5	1821.9	0.6	**34**	EIVHIQAGQCGNQIGAK
	653.8	3	1958.4	1958.0	0.4	**35**	GHYTEGAELVDSVLDVVR
	933.6	3	2797.8	2797.3	0.4	**33**	SGPFGQIFRPDNFVFGQS-GAGNNWAK

Continued on Next Page...

Table 5.6 – Continued

	Experimental			Calculated			
Name	**m/z**	**z**	**m**	**m**	**Δm**	**Score**	**Sequence**
TBB2C_HUMAN	566.0	2	1130.0	1129.6	0.4	**32**	FPGQLNADLR
Tubulin	572.5	2	1143.0	1142.6	0.4	**55**	LAVNMVPFPR
beta-2C chain	615.5	2	1229.0	1228.6	0.4	**87**	ISEQFTAMFR
Mass: 50255	660.5	2	1319.0	1318.7	0.3	**61**	IMNTFSVVPSPK
Score: 427	664.9	2	1327.8	1327.6	0.1	**62**	INVYYNEATGGK
	723.9	2	1445.8	1445.7	0.1	**58**	EVDEQMLNVQNK
	801.6	2	1601.2	1600.8	0.4	**33**	AVLVDLEPGTMDSVR
	846.6	2	1691.2	1690.9	0.3	**37**	ALTVPELTQQMFDAK
	608.5	3	1822.5	1821.9	0.6	**34**	EIVHLQAGQCGNQIGAK
	653.8	3	1958.4	1958.0	0.4	**35**	GHYTEGAELVDSVLDVVR
	933.6	3	2797.8	2797.3	0.4	**33**	SGPFGQIFRPDNFVFGQS-GAGNNWAK
TBB3_HUMAN	566.0	2	1130.0	1129.6	0.4	**32**	FPGQLNADLR
Tubulin	572.5	2	1143.0	1142.6	0.4	**55**	LAVNMVPFPR
beta-3 chain	615.5	2	1229.0	1228.6	0.4	**87**	ISEQFTAMFR
Mass: 50856	660.5	2	1319.0	1318.7	0.3	**61**	IMNTFSVVPSPK
Score: 302	808.6	2	1615.2	1614.8	0.4	**40**	AILVDLEPGTMDSVR
	846.6	2	1691.2	1690.9	0.3	**37**	ALTVPELTQQMFDAK
	608.5	3	1822.5	1821.9	0.6	**34**	EIVHIQAGQCGNQIGAK
	653.8	3	1958.4	1958.0	0.4	**35**	GHYTEGAELVDSVLDVVR
TBA1B_HUMAN	1015.5	1	1014.5	1014.6	-0.1	**11**	DVNAAIATIK
Tubulin	543.5	2	1085.0	1084.6	0.4	**50**	EIIDLVLDR
alpha-1B chain	625.4	2	1248.8	1248.5	0.2	**50**	YMACCLLYR
Mass: 50804	793.0	2	1584.0	1583.7	0.2	**42**	SIQFVDWCPTGFK
Score: 277	851.6	2	1701.2	1700.9	0.3	**46**	AVFVDLEPTVIDEVR
	573.7	3	1718.1	1717.9	0.2	**17**	NLDIERPTYTNLNR
	879.1	2	1756.2	1756.0	0.2	**47**	IHFPLATYAPVISAEK
	913.2	2	1824.4	1824.0	0.4	**38**	VGINYQPPTVVPGGDLAK
	933.0	2	1864.0	1863.9	0.1	**47**	AVCMLSNTTAIAEAWAR
	1004.5	2	2007.0	2006.9	0.1	**88**	TIGGGDDSFNTFFSETGAGK
TBA1A_HUMAN	1015.5	1	1014.5	1014.6	-0.1	**11**	DVNAAIATIK
Tubulin	543.5	2	1085.0	1084.6	0.4	**50**	EIIDLVLDR
alpha-1A chain	625.4	2	1248.8	1248.5	0.2	**50**	YMACCLLYR
Mass: 50788	800.0	2	1598.0	1597.8	0.2	**31**	TIQFVDWCPTGFK
Score: 263	851.6	2	1701.2	1700.9	0.3	**46**	AVFVDLEPTVIDEVR
	573.7	3	1718.1	1717.9	0.2	**17**	NLDIERPTYTNLNR
	879.1	2	1756.2	1756.0	0.2	**47**	IHFPLATYAPVISAEK
	913.2	2	1824.4	1824.0	0.4	**38**	VGINYQPPTVVPGGDLAK
	933.0	2	1864.0	1863.9	0.1	**47**	AVCMLSNTTAIAEAWAR
	1004.5	2	2007.0	2006.9	0.1	**88**	TIGGGDDSFNTFFSETGAGK

Continued on Next Page...

Table 5.6 – Continued

	Experimental			Calculated			
Name	**m/z**	**z**	**m**	**m**	**Δm**	**Score**	**Sequence**
TBA1C_HUMAN	1015.5	1	1014.5	1014.6	-0.1	**11**	DVNAAIATIK
Tubulin	543.5	2	1085.0	1084.6	0.4	**50**	EIIDLVLDR
alpha-1C chain	625.4	2	1248.8	1248.5	0.2	**50**	YMACCLLYR
Mass: 50548	800.0	2	1598.0	1597.8	0.2	**31**	TIQFVDWCPTGFK
Score: 244	851.6	2	1701.2	1700.9	0.3	**46**	AVFVDLEPTVIDEVR
	573.7	3	1718.1	1717.9	0.2	**17**	NLDIERPTYTNLNR
	879.1	2	1756.2	1756.0	0.2	**47**	IHFPLATYAPVISAEK
	913.2	2	1824.4	1824.0	0.4	**38**	VGINYQPPTVVPGGDLAK
	1004.5	2	2007.0	2006.9	0.1	**88**	TIGGGDDSFNTFFSETGAGK
	922.2	3	2763.6	2763.3	0.3	**8**	AYHEQLTVAEITNACFEPA-NQMVK

Surprisingly, ARNO was not identified. Actually, even in absence of proteome it was difficult to detect the ΔPBR signal. In fact, only in one of two experiments it was possible to unambiguously identify one ΔPBR peptide in the enriched sample (Table 5.5). Since the signal intensity was already particularly low, it was improbable to find ARNO signals in the proteome sample, where other (biotinylated) proteins are competing.

Furthermore, in the same experiment, an unexpected observation was made. In fact, labelling of proteome with SecinH3-TPD resulted in identification of various tubulin isotypes. For example, more than 10 peptide matches with good scores were found for "Tubulin beta chain" (Table 5.6). With a sequence coverage of 35 %, this hit can definitely be trusted. Thus, I decided to investigate the interaction of SecinH3-TPD with tubulin in more detail.

5.3.7 Investigation of the SecinH3/tubulin interaction

The identification of tubulin raised various questions. Is the interaction specific or simply an artefact due to the high concentration of tubulins in cells? Is the binding specific to SecinH3 or amenable to the photoreactive unit? And finally, is the interaction biologically relevant? To solve this questions, some informations about tubulin biology are needed.

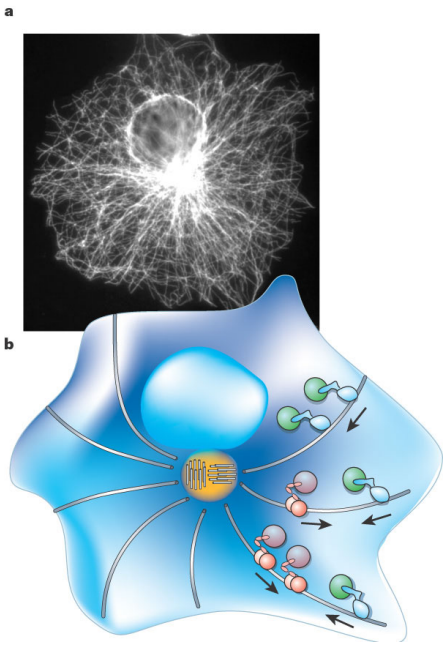

Figure 5.31: Microtubules in the cell - a. An interphase cell stained with an antibody to tubulin. Microtubules extend from the centrosome throughout the cell. **b.** A schematic diagram of the cell. Centrioles are shown in the centrosome (yellow). Red circles denote vesicles moving to the outside of the cell. Green circles denote vesicles moving to the centrosome.[137] *Reprinted by permission from Macmillan Publishers Ltd: Nature (Ref. 137), copyright 2003.*

5.3.7.1 Tubulin and the microtubules

Tubulin is the heterodimeric building block of microtubules and, as such, involved in various essential cell functions. Microtubules, as part of the cytoskeleton, offer mechanical support for the cell shape and provide tracks for transport of vesicle and organelles (Fig. 5.31). Moreover, the mitotic spindle they form during cell division is required for the correct segregation of chromosomes.

These various functions are highly regulated at different levels. Additionally to the transcription of different tubulin isotypes, post-translational modifications (PTMs) and interaction with microtubule-associated proteins (MAPs) further tune the nucleotide regulated microtubule dynamics[138].

Microtubules dynamic instability

By head-to-tail association of the α-β tubulin dimers, linear protofilaments are formed, which, in turn, lead to the formation of the cylindric microtubules by lateral association. The polarity of the dimer is maintained during the polymerisation process and mirrored in the microtubules polarity[137] (Fig. 5.32**a**). In animal cells, the minus end, which is terminated by the α subunit, is generally anchored at centrosomes and the plus

end grows toward the periphery, resulting in the characteristic radiating pattern[137,139] (Fig. 5.31).

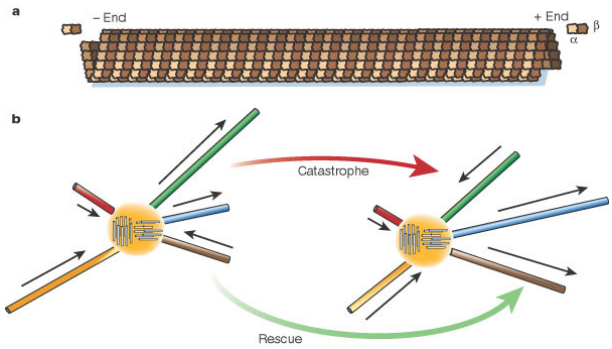

Figure 5.32: Microtubule structure and dynamics. - **a.** A microtubule lattice. The beta-subunit of tubulin is on the plus end. **b.** Dynamic instability of microtubules. Microtubules growing out from a centrosome switch between phases of growing and shrinking. The figure shows a hypothetical aster at two different times. The different colours represent different microtubules. The red and yellow microtubules are shrinking at both times. The blue microtubule is growing at both times. The green microtubule, growing at the first time, has undergone a catastrophe by the second time. The brown microtubule, shrinking at the first time, has undergone a rescue by the second time.[137] *Reprinted by permission from Macmillan Publishers Ltd: Nature (Ref. 137), copyright 2003.*

Microtubules polymerisation is a highly dynamic process. In fact, microtubules are subject to stochastic switch between growing and shrinking phases (Fig. 5.32**b**). This property, called dynamic instability[140], is essential to their function and depend on the GTPase activity of tubulin[141].

Indeed, each heterodimer subunit carries a guanine nucleotide. While GTP in the α unit is buried at the monomer-monomer interface within the dimer, and is thus non-exchangeable (N-site), in β tubulin it sits on the dimer surface and is fully exchangeable (E-site)[142]. For polymerisation to occur, the E-site has to be GTP-loaded, but, at the same time, binding of a new dimer triggers hydrolysis in the adjacent E-site. As a result, the microtubule body is composed of GDP-tubulin subunits and is in an energetically unfavorable state[141] which favours shrinking events (Fig. 5.33**b**). In fact, GDP-tubulin is thought to have a bent conformation which places stress on the microtubules lattice. During depolimerisation the ends curls, releasing the strain[137,141] (Fig. 5.33**d**).

An additional level of regulation is offered by MAPs. By interacting with the soluble, non-polymerized tubulin subunits, the microtubule wall lattice and/or microtubule

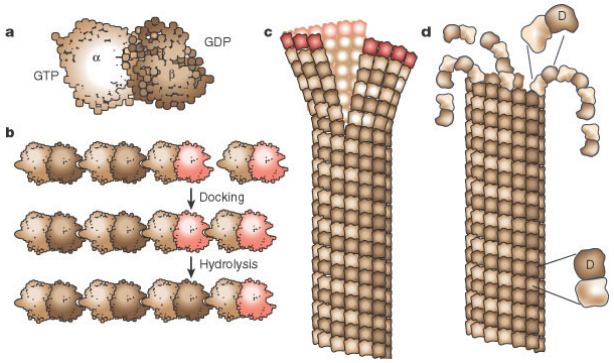

Figure 5.33: Model for how the GTP hydrolysis cycle is coupled to structural changes in the microtubule. - **a.** Atomic structure of the tubulin dimer as seen in the wall of the protofilament. **b.** Docking of the alpha-beta subunit to the microtubule end. Residues from the incoming alpha-subunit trigger hydrolysis of the GTP bound to the lattice-attached beta-subunit. **c, d.** Microtubules at growing ends contain sheets of protofilaments while microtubules at shrinking ends curl. The straight-bent transition is also shown in panel d. The GTP dimer is thought to have a straight conformation that fits nicely into the straight wall of the microtubule. Hydrolysis of GTP induces a bend in the subunit, but this bend is constrained within the lattice. The constraint places stress on the lattice, which is released during depolymerization, allowing the protofilament to adopt a curled conformation.[137] *Reprinted by permission from Macmillan Publishers Ltd: Nature (Ref. 137), copyright 2003.*

ends, they regulate microtubule stability and dynamics (See Refs. 143 and 144 for review).

When it comes to assigning a specific function to a microtubule, PTMs play a prominent role. In fact, a considerable number of reversible PTMs, such as acetylation, polyglycylation, polyglutamylation, tyrosination/detyrosination, phosphorylation, and palmitoylation contribute to microtubules diversity[145]. Most PTMs occur on microtubules rather than on unpolymerized tubulin and stable microtubules accumulate more modifications than dynamic microtubules[146]. Since differences in PTM patterns can be seen between stable microtubules, the PTMs are postulated to play a role in their specific functions (see Refs. 145 and 146 for review).

Microtubule-targeting drugs

Since during mitosis highly coordinated microtubule dynamics is required, compounds which interfere with microtubule dynamics limits proliferation and have been used as anticancer drugs for years. They are classified as microtubule stabilizers or destabilizers

but both classes inhibit mitosis through a similar mechanism of slowing microtubule dynamics, resulting in mitotic arrest and apoptosis[147].

An additional classification is done by the binding site on tubulin. Indeed, destabilising drugs are usually binding to the vinca- and colchicine-sites, while microtubule stabilisers mostly bind the taxane-site (but inhibitors binding at other sites were recently discovered). For a detailed explanation of the mechanisms of action, please read Ref. 147.

5.3.7.2 Labelling of purified tubulin

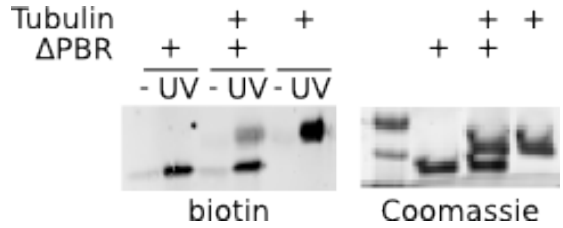

Figure 5.34: SecinH3-TPD binds ΔPBR preferentially over tubulin - Labelling of purified tubulin was tested in the standard labelling assay (1 µM protein, 20 µM SecinH3-TPD in 10 % DMSO). After separation on a 7.5 % SDS-PAGE and Western blotting, the labelled proteins were detected with a NeutrAvidin DyLight 800 fluorescent conjugate. The total amount of protein was detected by Coomassie staining of gel. The strong tagging of tubulin is competed by ΔPBR.

Because of the essential role of tubulin in cells, it was particularly important to investigate if SecinH3 is specifically binding to tubulin and interfering with its functions.

As a first control, purified tubulin was analysed in the standard labelling assay. As evident in Figure 5.34, SecinH3-TPD strongly labelled tubulin, but the binding was competed by ΔPBR. Indeed, while labelling of tubulin was drastically decreased in the presence of ΔPBR, ARNO signal remained stable. This result shows that SecinH3-TPD interacts with tubulin but, *in vitro*, SecinH3-TPD preferentially binds ARNO.

5.3.7.3 Competition with SecinH3

To exclude that SecinH3-TPD interaction with tubulin is mediated by the photoreactive or desthiobiotin moieties, a competition experiment was performed. Labelling was carried out in the presence of non derivatised SecinH3 in excess. The clear inhibition of

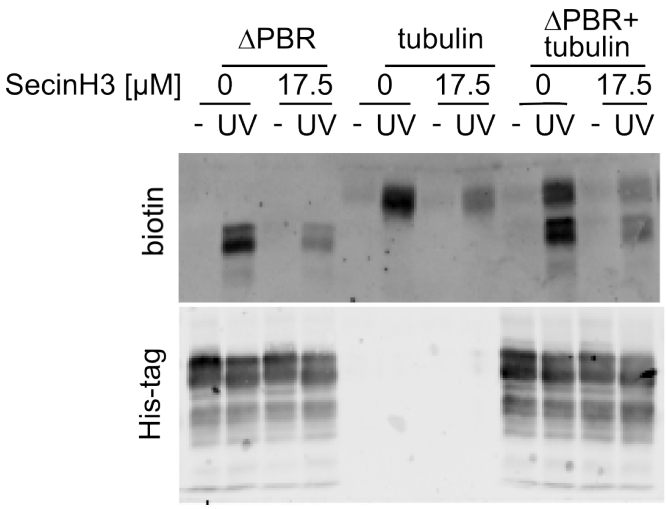

Figure 5.35: SecinH3 competes labelling of both tubulin and ΔPBR - Non derivatised SecinH3 was added in excess to labelling reactions with 2.5 μM SecinH3-TPD. After separation on a 7.5 % SDS-PAGE and Western blotting, the labelled proteins were detected with a NeutrAvidin DyLight 800 fluorescent conjugate. Tagging of both tubulin and ΔPBR was significantly reduced by the competitor.

labelling visible in Figure 5.35 in the competed sample, accords binding to the SecinH3 core.

5.3.7.4 Analysis of SecinH3 effect on microtubules structure

The most difficult question to answer was if the SecinH3-tubulin interaction is of biological relevance. To analyse a possible effect of SecinH3 on microtubule structure, immunofluorescence was used. H460 cells were treated overnight with SecinH3 or the microtubule-targeting drugs demecolcine and paclitaxel prior fixation in methanol. Microtubules were visualized by immunofluorescence using a FITC conjugated monoclonal anti-α-tubulin antibody.

In Figure 5.36, two representative views for each condition are shown. It is readily visible that mitosis is heavily impaired in the demecolcine and paclitaxel treated cells (these drugs binds the colchicine- and taxane-site, respectively). Additionally, in the lower paclitaxel panel, bundling of interphase microtubules, a typical effect of taxane[147,148], is apparent. In contrast, in SecinH3 treated cells the mitotic spindle is still formed (upper panel) and no grave defect is visible in the interphase cells. This results exclude a major impact of SecinH3 on microtubules dynamic. Possible smaller

effects, for example on PTMs or interaction with MAPs, can not be ruled out.

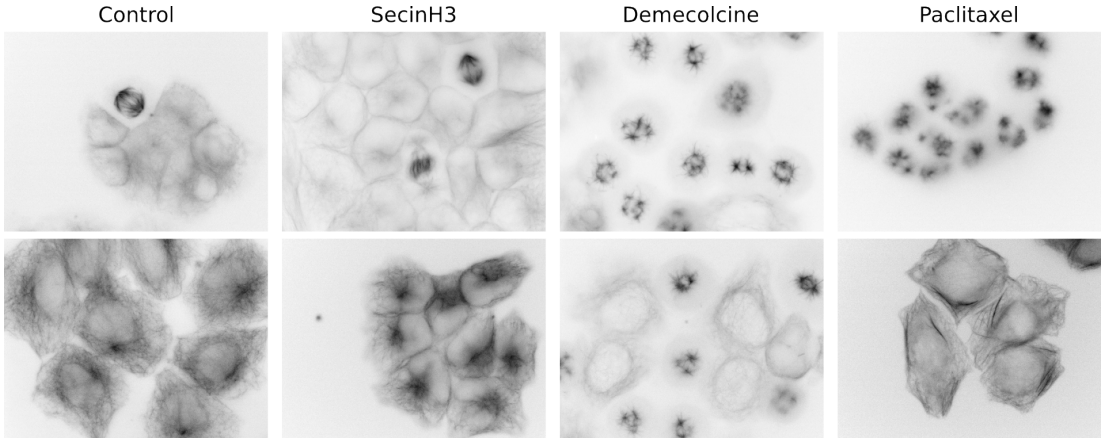

Figure 5.36: SecinH3 has no evident effect on microtubules structure - H460 cells were treated with SecinH3, demecolcine or paclitaxel, fixed in metanol and the microtubules structure visualized with a FITC conjugated monoclonal anti-α-tubulin antibody by immunofluorescence. In the demecolcine and paclitaxel treated cells, impairment of mitotic spindle formation is evident. The tipical bundling of interphase microtubules is evident for paclitaxel treated cells (lower panel). In contrast, no evident defect is visible in SecinH3 treated cells and the mitotic spindle is still formed (upper panel).

5.3.8 Discussion

The implementation of a CCMS protocol is a technically challenging task which requires separate optimisation of the single methods, as just described in this chapter. The first stage was the synthesis of the trifluorophenyl diazirine derivatised SecinH3 probe SecinH3-TPD. Because of the high costs of the TPD acid needed in the last step, the synthesis was first carried out with the non photoreactive 4-ethylbenzoic acid. This way, we could successfully adapt the procedure to reduce light exposition of the photoreactive moiety to a minimum.

At the stage of labelling of ARNO with the SecinH3 probe, solubility was the limiting factor. Indeed, Figure 5.11 (p. 78) shows that considering a SecinH3 maximal solubility of $\sim$15 µM (in 10 % DMSO) and an estimated K_D in the same range, a maximum of 50 % SecinH3/ARNO complex over total ARNO can be achieved. With the intent to improve SecinH3 solubility, diglyme was tested as an alternative solvent. Although solubility was improved (100 µM in 20 % diglyme), diglyme could not be used in the assay because it seemed to affect ARNO activity.

In fact, the solubility difficulties were not restricted to the compound. Figure 5.18 (p. 83) shows that under the labelling conditions most of ARNO was found in the pellet after centrifugation, indicating aggregation or coating of the protein. Addition of 0.005 % Triton X-100 to the buffer, allowed to recover ARNO back in solution. Under this conditions it was possible to detect labelling of ARNO by SecinH3-TPD and labelling of various mutants was tested.

Interestingly, ΔPBR, a construct lacking the polybasic region at the C-terminus, was more strongly labelled that the Sec7 domain and the full length ARNO was labelled similarly to the ΔPBR (Fig. 5.19). This is surprising because SecinH3 was identified using the Sec7 domain only of cytohesin-1[17] and, although selected on the full length protein, its parent aptamer M69 is a pure cytohesin Sec7-domain binder, since the same K_D was measured for the full length protein and the Sec7-domain alone[32]. There is thus no evident reason, why full length ARNO should bind SecinH3 better than ARNO-Sec7. Experiments like those described in Figures 5.14 (p. 80) and 5.34 (p. 102), where preference of SecinH3-TPD for ARNO over other proteins is clearly shown, exclude increased unspecific binding due to the bigger protein size. It should however be noted that variability in labelling in different purification charges of ΔPBR was noted and for the comparison a single charge of both constructs was used. It would therefore be worth to analyse if better labelling of ΔPBR is given for all purification batches.

Having shown that SecinH3-TPD retains its binding capacity and is able to label ARNO, the next step was the analysis of the binding site by MS. As the expertise for analysis of proteins and peptides by MS was lacking in our group, the reference peptide synthesised for the establishment of the measuring protocol offered fundamental help. Indeed, it did not only allowed successful optimisation of sample preparation and analysis conditions but it was also essential in fixing and setting the MS instrument.

Analysis by LC-MS was a good training, but the use of tandem LC-MS emerged as necessary for unambiguous identification of the peptides. Indeed, tryptic digestion of ARNO gives various peptides with similar masses which can not be discerned without some sequence data. Moreover, analysis of LC-MS/MS data by specific software was easier and allowed identification of unknown proteins. This was particularly important in the second part of this study: the analysis of SecinH3 specificity in proteome.

The amount of labelling of ARNO by SecinH3-TPD turned out not to be sufficient for identification of the modified peptide. Indeed, in contrast with the analysis of

the almost 100 % modification of cysteines with iodoacetamide (Fig. 5.27, p. 90), no difference could be detected between the chromatograms of labelled and non labelled ARNO. Possible explanations are that SecinH3-TPD labels either multiple peptides nearby the binding site (and thus leads to a complex distribution of various modified peptides in the sample) or one of the few peptides which are not identified under the analysis conditions. However, it is more probable that the amount of labelled protein is simply too low with respect to the total protein to allow its detection in a complex sample.

Enrichment of the modified proteins or fragments with streptavidine before or after digestion, respectively, could allow their detection even with low labelling yield. Unfortunately, experiments in that direction where not successful. Indeed, although prebiotinylated ARNO could be selectively enriched with streptavidine-beads, digested on beads and identified in MS (Tab. 5.2 and 5.3, p. 92), it was not possible to detected the biotinylated peptide after elution from the beads. It was not possible to define if this was because of poor elution from the beads, coelution of contaminants disturbing analysis or perhaps degradation of the petides during elution. Application of the same protocol to SecinH3-labelled ARNO lead to identification of a single relevant peptide with Mascot (Tab. 5.5, p. 96). The signal was very weak and thus also in this case the modified peptide could not be identified.

Yet, the fact that not only prebiotinylated but also SecinH3-labelled ARNO was identified after the whole CCMS procedure confirms that the protocol was successfully established. While prebiotinylated ARNO could be enriched and identified also out of complex samples (i.e. cell lysate) the same was not possible for labelling reactions in cell lysate, even when purified ARNO was added to the sample before irradiation. Considering the weak signal in the analysis above with purified ARNO, the result is not surprising. Indeed, in cell lysate both the labelling reaction and the enrichment are complicated by the high amount of protein material present in the sample. Here, again, more efficient labelling could significantly improve the results, but this can only be achieved with higher concentrations or a better binding probe. In fact, the use of photoreactive affinity probes is mostly reported for inhibitors with affinities in the nM range[87,93,134]. For less tight interactions, high concentrations and excess of probe (up to low mM concentrations) were used[86,136].

Western blot analysis of labelling in cell lysate, demonstrated that SecinH3-TPD is able to label ARNO even in complex samples and competition experiments showed that labelling can be reduced by the addition of unmodified SecinH3 (Fig. 5.30). Interestingly, disruption of the SecinH3-TPD/ARNO complex favoured labelling of (an) other unknown protein(s) (Fig. 5.30**b**, lane 16). Since the signal at ∼80 kDa is much stronger in the presence of SecinH3, the interaction can not be ascribed to the SecinH3 core but rather to the photoreactive or reporter groups. The fact that at least three biotin-binding proteins around 80 kDa in size are known (Tab. 5.4, p. 95), seems to support this hypothesis which, however, was not tested so far. CCMS analysis of a sample labelled in the presence of SecinH3 could reveal the identity of the protein. However, should the binding to desthiobiotin be very tight, this could complicate the enrichment step and possibly prevent identification of the protein.

The identification of tubulins by CCMS experiments in cell lysate (Tab. 5.6, p. 96) is an additional proof of the successful implementation of this method. Indeed, this result shows, that selectively labelled protein can be identified out of cell lysate with my protocol. Interaction of tubulin and SecinH3 was confirmed *in vitro*, although SecinH3-TPD was shown to preferentially bind ARNO (Fig. 5.34, p. 102) and competition with non derivatised SecinH3 excluded that the interaction was mediated by the photoreactive moiety or the photoreactive tag (Fig. 5.35, p. 103).

Still, this finding poses some questions. First of all: why only tubulin is found if affinity for ARNO is higher? One could argument that the purified ARNO used for the comparison is actually a recombinantly expressed protein and could thus have a slightly different folding as the native protein in cell lysate. However, samples spiked with equal concentration of recombinant ARNO neither allowed its identification. Furthermore, it was shown that the lysis buffer does not derange the labelling reaction (Fig. 5.29, p. 93). It could also be, that tubulin is actually not labelled stronger than ARNO in cell lysate but it is simply better digested by trypsin or the petides are more efficiently ionised during MS. Both cases would lead to higher signals for tubulin but, in fact, MS analysis of simple samples of both protein gave comparable results. That ARNO can be actually detected by CCMS was shown with prebiotinylated ARNO, excluding loss of the peptides during work up due to low solubility or other effects. Thus, at this stage I can not offer any conclusive explanation for the identification of tubulin instead of ARNO.

A second question concerns the biological relevance of the tubulin/SecinH3 interaction: does binding of SecinH3 to tubulin affect any cellular process? As described in Section 5.3.7.1 (p. 99), the microtubules, and thus their building blocks tubulins, play an important role in the control of cell function and morphology. Their involvement in the most disparate processes like mitosis and exo/endocytosis makes answering this question quite laborious, but a first help comes from the SecinH3 literature. Microtubules are known to be essential for maintaining structure and function of the Golgi complex: although the effect of microtubule-targeting drugs is less drastic than that of, e.g., brefeldin A, depolymerization of microtubules leads to separation and redistribution of the Golgi subcompartments[139]. However, already the first publication about SecinH3 showed that no significant disturbance of Golgi integrity is observed upon SecinH3 treatment[17]. Moreover, Figure 5.36 (p. 104) shows that, in contrast to the microtubule-targeting drugs demecolcine and paclitaxel, SecinH3 treatment had no dramatic effect of microtubules structure. So far, thus, nothing indicates a major effect of SecinH3 on microtubules function. Possible minor effects, for example at the level of post-translational modifications or interaction with microtubule-associated proteins, can not be excluded at this stage but their investigation is far beyond the scope of this project.

6

Conclusions

During this PhD project, very diverse methods for interaction analysis were implemented and applied to the investigation of the activity mechanisms of cytohesins and their small molecule inhibitors. Cytohesin were chosen because the discovery of their dual function as both guanine nucleotide exchange factors (GEFs) and ErbB receptor activators raised the question on how these relatively small proteins regulate their different roles.

The main focus of this work was on the establishment of the new methodologies. These are of particular interest for a group like ours, where high throughput screenings deliver not only new small molecules but from time to time also unexpected new target proteins.

6.1 BRET

The first technique which had to be established was a cellular BRET assay. In a first approach, a possible interaction between EGFR and ARNO was tested by analysis of energy transfer from EGFR-Luc to ARNO-GFP. The observation that C-terminal GFP labelling affects ARNO functionality (A. Bill, J. Theis, unpublished results) explains why no binding was detected, although we showed later on by a different method, that ARNO directly interacts with the EGFR[50].

Since in the meanwhile the question if ARNO interacts with the EGFR was answered, BRET with N-terminally labelled ARNO, which had already been shown to be active[29–31,53–59], was not tested. This experiment could be of interest only if additional

analyses of the EGFR-ARNO interaction are planned. For example, a screening for inhibitors of this interaction could be envisaged: the assay would have the advantage over standard *in vitro* assays of selecting only inhibitors which are active in cell culture. The generation of a cell line which stably expresses EGFR-Luc and ARNO-GFP would be necessary for high throughput screening and probably solve some of the problems encountered during analysis with cells transiently expressing the two proteins.

Although the analyses above did not deliver an answer to the biological question if ARNO interacts with the EGFR, I could establish a set of control proteins and experiments which proved important in the subsequent measurements. Moreover, some important defects of the Mithras LB 940 plate reader used were identified and adjusted.

In the second BRET approach, the possible interaction between EGFR and ARNO was observed indirectly by monitoring the evolution of energy transfer from EGFR-Luc to EGFR-GFP upon coexpression of unlabelled ARNO. Starting from an already relatively high BRET-signal, expression of ARNO increased further the measured BRET-ratio (Fig. 3.4, p. 21). An increase in signal can be explained by either an induction of dimerisation or a conformational change of the receptors (clustering had at that time already been excluded by superresolution light microscopy experiments[50]).

To verify the BRET system, I analysed the effect of EGF stimulation on the BRET-ratio of the EGF receptors. Since upon stimulation important changes happens both in the conformation and the amount of dimerised receptors, a change in BRET-signal was expected. Yet, no difference was detected.

Later on, I observed that our EGFR constructs, which lacked the exon 4 in the extracellular EGF binding domain, did actually not respond to stimulation (Fig. 3.8, p. 23). Accordingly, Wang *et al.* reported very recently the identification of an exon 4-deletion variant of the EGFR which displayed only minimal EGF binding activity and underwent ligand-independent autophosphorylation and self-dimerisation[52].

Given the observations above, no change in BRET-signal could be expected upon stimulation. Moreover, since the ARNO overexpression experiments were performed with dimeric EGFR, the experiments above shows that ARNO acts on already dimerised receptor. That A. Bill demonstrated the same with anisotropy microscopy experiments[50], can be taken as a validation of the results.

Still, it should be stressed out that the established BRET system was not optimal. In particular, the very low signals detected and the limitations due to transfection made

the measured BRET-ratio quite variable. Especially the analyses with overexpressed ARNO would need further validation, in particular with respect to receptor expression.

Some stimulation experiments were also performed with full length receptor constructs. However the variation in luminescence and fluorescence signal was too high to allow analysis of the acquired data. Particularly problematic was the use of DeepBlueC as a substrate for the luciferase, since due to the rapid decays of its luminescence, it is not possible to measure more than one time point per sample. Moreover, DeepBlueC itself decayed rapidly in the injection system making time-dependent measurement very difficult. The use of Coelenterazine h or EnduRen as substrates and the respective fluorophores (BRET[1] and eBRET combinations), which allow detection for up to one to several hours[46], would have been better suited. Since Coelenterazine h has over 100-fold higher quantum yield than DeepBlueC, its use could have solved also the problem of low luminescence.

Looking back at the ensemble of the BRET experiments it is clear that the initial choice of the protein constructs and the BRET combination was unpropitious. Adopting the BRET plasmids and system that were already partially established at the start of my thesis was an unfavourable decision, which I should have questioned earlier in the project. Nevertheless some promising results could be recorded.

6.2 SPR

The SPR analysis of the EGFR/ARNO interaction lead to suboptimal binding data due to complex responses in all the tested conditions. However, the intimacy developed with the Biacore machine was of undeniable value for the subsequent experiments. Having already seen a comprehensive set of possible problems, allowed to quickly recognise and correct technical irregularities.

The measurements on immobilised GST-EGFR provided promising data. With the knowledge of the subsequent experiments, it is now possible to suggest measuring parameters for clean binding curves. Low concentrations of ARNO ($< 1\,\mu M$) and new coupling reagents should be used. If the K_D should be higher than $1\,\mu M$, it is worth trying the inverse setup with coupled ARNO to see if the EGFR can be used in higher concentration. This would allow a preciser determination of the binding parameters.

In the second measurements group, we observed binding of ARNO-Sec7 to immobilised second generation cytohesin inhibitor Secin16 with a dissociation constant of ∼150 nM (Fig. 4.24, p. 57). The data was reproducible and measurements at different surface densities lead to comparable results. Note that high Sec7 concentrations of above 500 nM lead to complex binding behaviour and should therefore not be taken into consideration.

Positively, no binding was detected to the negative control compound XH1009 (Fig. 4.26, p. 59). Since GST turned out to be an unsuitable negative control, some more control proteins are currently being measured. Preliminary data seems to confirm the specificity of the measured interaction.

Unfortunately, it was not possible to determine the binding parameters to SecinH3 because of unpredictable behaviour of the SecinH3 derivatised surface. The use of SecinH3 with propanolamine instead of ethanolamine linker could possibly solve this problem. Indeed, both Secin16 and XH1009 had this (longer) linker. SecinH3 binding data would permit direct comparison of the affinities of the new inhibitors with that of the parent molecule.

In conclusion, an SPR method for the analysis of small molecule/protein interaction was successfully implemented and delivered meaningful kinetic parameters. However, I would only recommend this method if kinetic parameters are needed and not if only a quick answer about binding/non binding is desired. Indeed, the system has to be fine tuned in an extremely time-consuming procedure not only for every protein, but also for the single compounds . Moreover, with the actual setup, every inhibitor has to be derivatised and its activity confirmed separately. Since the behaviour of the compound on the chip cannot be predicted, coupling of many derivatisations of a single compound can be necessary and would increase drastically the costs of analysis. SPR measurements are also particularly buffer dependent, especially with regard of the protein. In absence of an adequate positive control (which is mostly not available at the early stage of drug discovery), the identification of the right buffer could prove very challenging. For all these reasons I recommend using SPR only if kinetic data are needed and only for a limited amount of inhibitors.

A positive feature of this system was the stability of the derivatised chips. In fact, we were able to reproduce binding experiments on the same chip after 3 months of storage

at 4°C. This could be particularly interesting as a routine test of protein activity, since every new purified charge could be easily tested by looking at the binding affinity.

6.3 CCMS

The first stage of the CCMS protocol was the synthesis of the photoreactive probe SecinH3-TPD. The coupling of the commercially available diazirine benzoic acid via the active benzoyl chloride to **BA99**, as done in the synthesis of SecinH3-TPD (Fig. 5.9, p. 75), can in principle be applied to any other compound where a primary amine is available. For example, since changes at the *tert*-butyl substituent of Secin16 were shown not to influence its activity (J. S. Hannam, unpublished), the reaction in Figure 6.1 could be envisaged for introduction of the photoreactive diazirine moiety in Secin16.

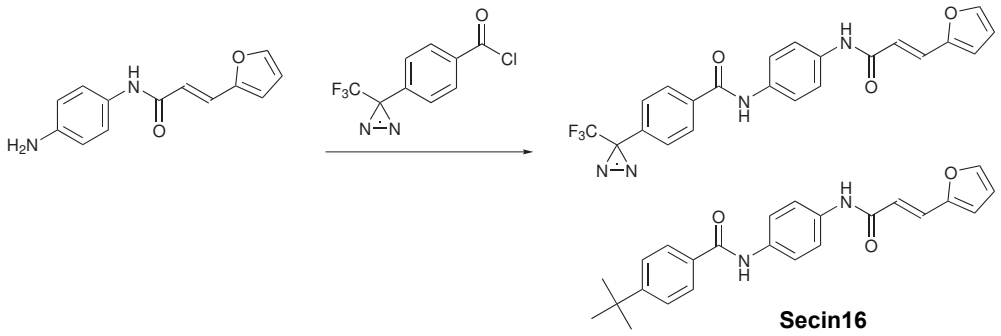

Figure 6.1: Possible reaction for the introduction of the diazirine moiety in Secin16 - Coupling of the commercially available diazirine benzoic acid via the active benzoyl chloride is in principle possible for every primary amine containing molecule.

SecinH3-TPD was shown to selectively label ARNO *in vitro* (Fig. 5.14, p. 80). The addition of 0.005 % Triton X-100 was necessary to keep ARNO in solution and did not disturb the labelling reaction (Figs. 5.18 and 5.19, p. 83). Unfortunately, the labelling yield was limited by the low solubility and relatively high K_D of SecinH3 and was too low to allow identification of the labelling site by MS. In an analysis where the cysteines of ARNO where modified with iodoacetamide, I was able to detect the modification both in the complete protein and in the digested sample (Fig. 5.28, p. 91). This was not the case for ARNO labelled with SecinH3-TPD. Of course, alkylation was a best case sample, since the reaction had almost 100 % yield and only three possible reaction sites. In the case of photoreactive labelling the conditions are more complex. Indeed, the

photoreactive group can covalently insert into any amino acid in the binding region. If these are distributed on more than a single tryptic peptide a mixture of labelled peptides is obtained. When additionally the labelling yield is strongly limited by solubility, the amount of every single modified peptide could be below the detection limit. Since the solubility of SecinH3 could not be increased, the only possibility to obtain a simpler sample would be to use an enzyme which cuts less frequently than trypsine for the digestion (e.g. LysC). The longer fragments could reduce the complexity of the sample and lead to less modified samples in higher amounts. However, the separation step could become more challenging.

By Western blot analysis of labelling in cell lysate I showed that SecinH3-TPD labels ARNO even in complex samples and competition experiments revealed that labelling is reduced by the addition of unmodified SecinH3 (Fig. 5.30, p. 94) and could thus be used for the analysis of specificity.

The CCMS protocol for the analysis of binding specificity consists in labelling in cell lysate and enrichment of the labelled proteins via the reporter group before digestion, LC-MS analysis and data base search. The procedure was optimised with success on prebiotinylated ARNO added to cell lysate (Tab. 5.2, p. 92). Concentrations down to 50 nM bio-ΔPBR lead to identification of ARNO by Mascot search and the lysate was shown not to interfere with the enrichment (Tab. 5.3).

Application of CCMS to a ΔPBR labelling reaction identified a single ARNO fragment (Tab. 5.5, p. 96). The very low signal measured for this fragment was not present when the labelling was performed in cell lysate, even when purified ΔPBR was added before irradiation. Instead, tubulins, the building monomers of microtubules, were enriched in the SecinH3-TPD labelled proteome sample (Tab. 5.6, p. 96).

Interaction of SecinH3 with tubulins was confirmed by labelling *in vitro*. SecinH3-TPD labelled the purified tubulins and binding was competed by unmodified SecinH3. Still, SecinH3-TPD preferentially labelled ARNO (Figs. 5.34 and 5.35, p. 102). It is not clear why only tubulins are identified, although ARNO is, at least *in vitro*, the preferred binding partner of SecinH3-TPD.

The discovery that SecinH3 binds to tubulins lead to the question if SecinH3 influences the microtubules activity. By immunofluorescence with FITC conjugated monoclonal anti-α-tubulin antibody, I showed that SecinH3 does not seriously affect the

microtubules organisation. The microtubule-targeting drugs demecolcine and paclitaxel were used as control (Fig. 5.36, p. 104). Moreover, it is known that inhibition of microtubules leads to separation and redistribution of the Golgi subcompartments[139]. Yet, Hafner *et al.* showed that SecinH3 treatment does not affect Golgi integrity[17]. Thus, I can exclude a major effect of SecinH3 on microtubules. Still, minor effects on interaction with microtubule-associated proteins or post-translational modifications are possible. Their investigation, however, is far beyond the scope of this project.

6.4 Solubility

All the methods implemented were tested on ARNO and its small molecule inhibitors SecinH3 and Secin16. Now I know that these are by far not optimal testing samples, as they share a disadvantageous quality: bad solubility. The problem was particularly evident in the *in vitro* analysis.

In CCMS ARNO could be kept in solution by the addition of 0.005 % Triton X-100 to the buffers. Since relatively diluted solution were used, ARNO solubility was not particularly problematic in this assay. However the issue was evident in the Biacore measurements. Biacore depends on high quality reagents and even partial aggregation of the analyte is often detectable in the sensorgrams. Poor solubility of ARNO could account for the complex binding behaviour detected in both setups. The upper concentration limit I had to define, affected the quality of the measurements, since for the determination of precise binding parameters measurements at concentrations above the K_D are needed. All Biacore buffers contained 0.005 % Tween-20 as detergent. So far, we have not tested if Triton X-100 can improve the measurements.

With regard to the compounds, the situation is even more difficult. SecinH3 has a K_D of around 15 µM for ARNO and solubility of 15-20 µM. Since the K_D is by definition the concentration at which half of the protein is in complex with the compound*, there is no chance to have more than ∼50 % of ARNO bound to SecinH3. This is of course a difficult starting point for assays like CCMS, where the interest is in labelling as much ARNO as possible to improve detection, as well as for activity and cellular assays

*With $AB \rightleftharpoons A + B$, then $K_D = \frac{[A]\,[B]}{[A\,B]}$. If $[B] = K_D$, then $[A] = [A\,B]$.

The second generation cytohesin inhibitor Secin16 has better IC_{50} and K_D values. However, this improvement is countered by diminished solubility. Thus, its use would not help improving the labelling efficiency.

The solubility problem accompanied most of the experiments with the cytohesin inhibitors in the last years. The efforts up to now were mainly concentrated on improving the activity. However, a compound with similar activity but marked better solubility would be of much higher value for our research. Thus the directed insertion of solubility-improving modifications should be pursued.

6.5 Outlook

In this thesis three methods for binding analysis were presented. The BRET assay already require a relatively high amount of information on the mechanism and/or targets of the interaction to be investigated to be successfully implemented. Instead SPR and CCMS can be applied at earlier stage of hit characterisation after an high throughput screening. These last two methods are complementary in the information they deliver. After identification of the interaction partner(s) by CCMS, the binding characteristics could be determined by SPR. As both methods require modification of the inhibitor, it would be useful to have a common derivatisation for both systems. This would reduce the synthetic efforts and the time spent in activity analysis of the derivatised compound. An interesting step in this direction would be the use of "click" chemistry.

"Click" chemistry is a general term introduced by Sharpless[149] to describe chemical reactions which can be efficiently used for modular synthesis, have very high chemical yields and selectivity and can be preferably performed in water. The best known and mostly used "click" reaction is the Cu(I)-catalyzed Huisgen 1,3-dipolar cycloaddition between azides and alkynes.

The use of "clickable" probes has already been reported for CCMS where alkynyl ABPP probes are used for proteome labelling and the azide bound reporter group is added only afterwards[90,135] (Fig. 6.2). This strategy often offers improved specificity as the bulky reporter group is not present during labelling.

The same "click" strategy could be used to immobilise the small molecule on a "clickable" SPR chip. A quick literature survey revealed that the preparation of a "click-

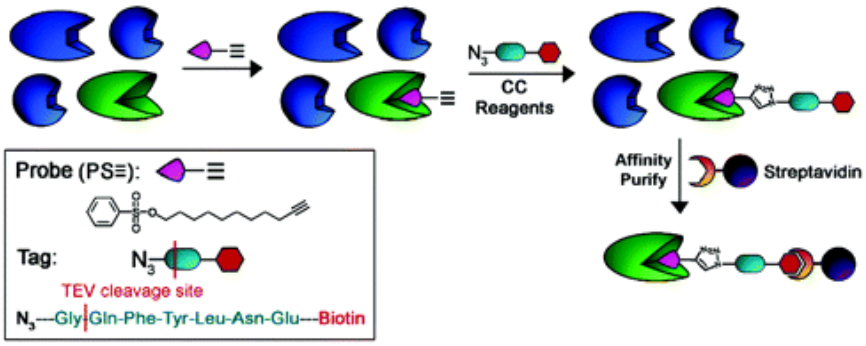

Figure 6.2: Example of "click" chemistry in CCMS - The labelling reaction is performed with an alkynyl probe before the reporter group is added by "click" chemistry. *Modified with permission from JACS (Ref. 90). Copyright 2005 American Chemical Society.*

able" SPR surface has already been reported for *"Clicked" carbohydrate self-assembled monolayers*[150] and should thus be feasible.

Moreover one could use "clickable" fluorophores, which are commercially available, to rapidly obtain fluorescently labelled inhibitors. These could be used in further assays like microscale thermophoresis[151] or fluorescence polarisation. The use of "clickable" inhibitors would thus allow to rapidly switch from an assay to another and reduces the time spent in the development of new synthetic routes.

7

Materials and methods

7.1 Methods

7.1.1 Organic synthesis

Chemicals were purchased from *Alfa Aesar*, *Aldrich*, *Bachem* and *Fluka* and used without further purification. For synthesis, solvents were obtained from *Fluka*, *Merck* and *Riedel de Haen* in 'per analysis' grade.

For column chromatography, solvents of technical grade were distilled before use and silica gel from *Acros*, with a particle dimension of 35-70 µm, was utilized. Thin layer chromatography was performed on aluminium plates of silica gel 60 F_{254} from *Merck*. Detection was done with ultraviolet (UV) light at 254 nm.

Nuclear magnetic resonance (NMR) spectra were measured on *Bruker AM 400* (^{1}H: 400 MHz, ^{13}C: 100 MHz) and *Bruker AM 300* (^{1}H: 300 MHz) spectrometers at room temperature. The chemical shifts are given in ppm relative to tetramethylsilane and the spin multiplicity as s = singulet, d = doublet, t = triplet, br = broad, m = undefined multiplet.

Mass spectrometry (MS) measurements were done on a *Bruker esquire HCT* spectrometer with atmospheric pressure interface-electrospray ionisation (API-ESI).

7.1.1.1 Synthesis of SecinH3-TPD

(2-(2-[5-Benzo[1,3]dioxol-5-yl-1-(4-nitro-phenyl)-1H-[1,2,4]triazol-3-yloxy]-ethoxy)-ethyl)-carbamic acid tert-butyl ester (BA91)

BS-1 (2-(2-(Benzo[1,3]dioxole-5-carbonylthiocarbamoyloxy)-ethoxy)-ethyl)-carbamic acid tert-butyl ester, 1.6 g, 3.9 mmol, 1.0 eq.) and 4-nitrophenylhydrazine (0.7 g, 4.74 mmol, 1.2 eq.) were suspended in EtOH (22 ml) and stirred at 80°C under reflux for 8 h. The reaction mixture was evaporated under vacuum and purified by column chromatography on silica gel (EtOAc/cyclohexane 100:0 – 45:55 gradient) to give the desired product (1.2 g, 58 %). Rf (EtOAc/cyclohexane 1:1): 0.83. ^{1}H NMR ($CDCl_3$): δ ppm 1.46 (s,

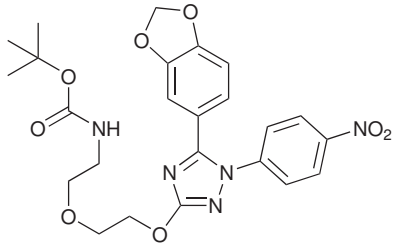

Figure 7.1: BA91 - Chemical formula: $C_{19}H_{19}N_5O_6$. Molecular weight: 413.4.

9 H) 1.87 (br. s., 2 H) 3.36 (t, J=4.53 Hz, 2 H) 3.63 (t, J=5.04 Hz, 2 H) 3.84 - 3.90 (m, 2 H) 4.50 - 4.56 (m, 2 H) 4.96 (br. s., 1 H) 6.06 (s, 2 H) 6.82 (d, J=7.81 Hz, 1 H) 6.95 (s, 1 H) 6.96 - 7.00 (m, 1 H) 7.57 (d, J=9.06 Hz, 2 H) 8.28 (d, J=8.81 Hz, 2 H).

2-2-[5-Benzo[1,3]dioxol-5-yl-1-(4-nitro-phenyl)-1*H*-[1,2,4] triazol-3-yloxy]-ethoxy-ethylamine (BA94)

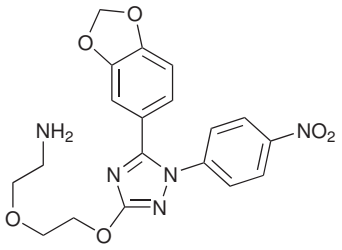

Figure 7.2: BA94 - Chemical formula: $C_{19}H_{19}N_5O_6$. Molecular weight: 413.38.

BA91 (800 mg, 1.6 mmol, 1 eq.) was dissolved in formic acid (14.5 ml) and stirred at room temperature for 2.5 h. The reaction was evaporated under vacuum, redissolved in formic acid (4 ml) and evaporated again. The residue was co-evaporated twice with toluene. Purification by column chromatography on silica gel (25 V_C MeOH/CH_2Cl_2 1:99, each 5 V_C MeOH/CH_2Cl_2/Triethylamine 1:98:1 – 3:96:1) gave the desired product (433 mg, 67 %). ^{1}H NMR ($CDCl_3$): δ ppm 3.17 (br. s., 2 H) 3.80 (br. s., 2 H) 3.86 (br. s., 2 H) 4.47 (br. s., 2 H) 6.01 (s, 2 H) 6.77 (d, J=8.06 Hz, 1 H) 6.89 (s, 1 H) 6.89 - 6.94 (m, 1 H) 6.99 - 7.41 (m, 4 H) 7.27 - 7.30 (m, 1 H) 7.53 (d, J=8.81 Hz, 2 H) 8.21 (d, J=9.06 Hz, 2 H).

6-(5-Methyl-2-oxo-imidazolidin-4-yl)-hexanoic acid (2-(2-[5-benzo[1,3]dioxol-5-yl-1-(4-nitro-phenyl)-1*H*-[1,2,4]triazol-3-yloxy] -ethoxy)-ethyl)-amide (BA96)

N-Ethyldiisopropylamine (830 µl , 4.8 mmol, 5.1 eq.) was added to a solution of HBTU (359 mg, 0.9 mmol, 1 eq.) and desthiobiotin (223 mg, 1.0 mmol, 1.1 eq.) in DMF (4 ml). The mixture was stirred for 10 min at room temperature. A solution of **BA94** in DMF (10 ml) was added dropwise and the reaction stirred for 1 h. The reaction mixture was partitioned between brine and CH_2Cl_2. The organic layer was dried over $MgSO_4$

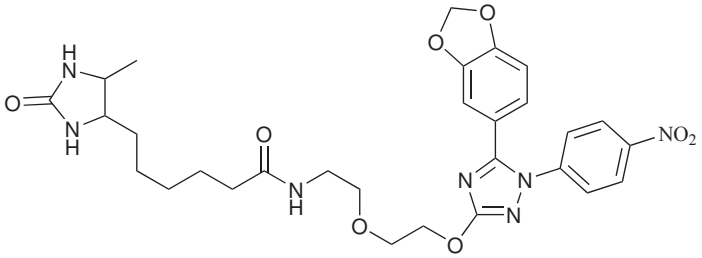

Figure 7.3: BA96 - Chemical formula: $C_{29}H_{35}N_7O_8$. Molecular weight: 609.63.

and evaporated. The resulting residue was purified by column chromatography (9 V_C MeOH/CH_2Cl_2 2:98, 4 V_C MeOH/CH_2Cl_2 10:90) to give the desired product (698 mg, quantitative) ^{1}H NMR ($CDCl_3$): δ ppm 1.12 (d, J=6.55 Hz, 3 H) 1.18 - 1.54 (m, 9 H) 1.66 (quin, J=6.86 Hz, 2 H) 2.21 (t, J=7.30 Hz, 2 H) 3.48 (q, J=5.04 Hz, 2 H) 3.65 (t, J=5.04 Hz, 2 H) 3.68 - 3.74 (m, 1 H) 3.80 - 3.85 (m, 1 H) 3.85 - 3.89 (m, 2 H) 4.50 - 4.54 (m, 2 H) 6.05 (s, 2 H) 6.51 (t, J=4.78 Hz, 1 H) 6.82 (d, J=7.81 Hz, 1 H) 6.94 (s, 1 H) 6.96 (d, J=1.26 Hz, 1 H) 7.56 (d, J=9.06 Hz, 2 H) 8.27 (d, J=8.81 Hz, 2 H).

6-(5-Methyl-2-oxo-imidazolidin-4-yl)-hexanoic acid (2-2-[1-(4-amino-phenyl)-5-benzo[1,3]dioxol-5-yl-1*H*-[1,2,4]triazol-3-yloxy] -ethoxy -ethyl)-amide (BA99)

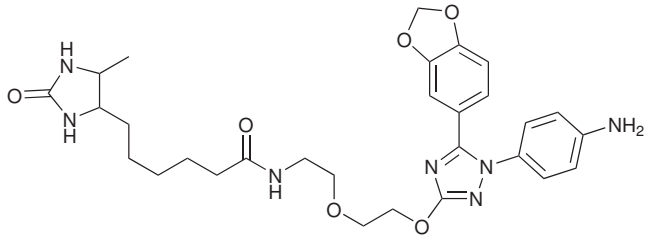

Figure 7.4: BA99 - Chemical formula: $C_{29}H_{37}N_7O_6$ Molecular weight: 579.6.

BA96 (530 mg, 0.9 mmol) was mixed with 10 % palladium on carbon (126 mg), THF (10 ml) and EtOH (10 ml) and added to a stirred autoclave. The reactor was purged with argon and pressurised with hydrogen (9 bar). After 2 h the reaction was filtered over celite and evaporated under vacuum. Purification by column chromatography on silica gel (MeOH/CH_2Cl_2 2:98 – 5:95 gradient) gave the desired product (235 mg, 50 %). Rf (MeOH/CH_2Cl_2/triethylamine 10:89:1) 0.66.

N-(4-[5-Benzo[1,3]dioxol-5-yl-3-(2-(2-[6-(5-methyl-2-oxo-imidazolidin-4-yl)-hexanoylamino]-ethoxy)-ethoxy)-[1,2,4]triazol-1-yl] -phenyl)-4-ethyl-benzamide (BA103)

4-Ethylbenzoic acid (16.3 mg, 0.1 mmol, 1 eq.) was dissolved in THF (225 µl), and 1 droplet of DMF added. The reactor was flushed with argon and a solution of oxalylchloride (69 µl, 0.5 mmol, 5 eq.) in THF (225 µl) added dropwise. The reaction

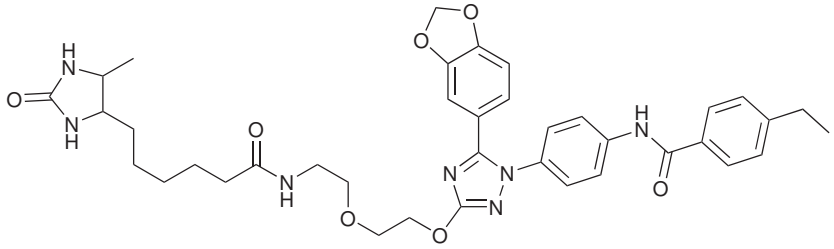

Figure 7.5: BA103 - Chemical formula: $C_{38}H_{45}N_7O_7$. Molecular weight: 711.81

mixture was stirred for 2.5 h at room temperature and the solvent evaporated under vacuum. The oxalyl chloride was removed by coevaporation with once THF and twice toluene. The crude product was dissolved in CH_2Cl_2 (500 µl) and the reactor purged with argon. **BA99** (64.6 mg, 0.1 mmol, 1 eq.) and *N*-Ethyldiisopropylamine (38 µl, 0.2 mmol, 2 eq.) in CH_2Cl_2 (1.8 ml) were added dropwise and the reaction stirred for 2 h. The reaction mixture was concentrated under reduced pressure. Purification on TLC plates (MeOH/CH_2Cl_2 10:90) gave the desired product (22 mg, 28 %).

N-4-[5-Benzo[1,3]dioxol-5-yl-3-(2-(2-[6-(5-methyl-2-oxo-imidazolidin-4-yl)-hexanoylamino]-ethoxy-ethoxy)-[1,2,4]triazol-1-yl] -phenyl)-4-(3-trifluoro-methyl-3*H*-diazirin-3-yl)-benzamide (SecinH3-TPD)

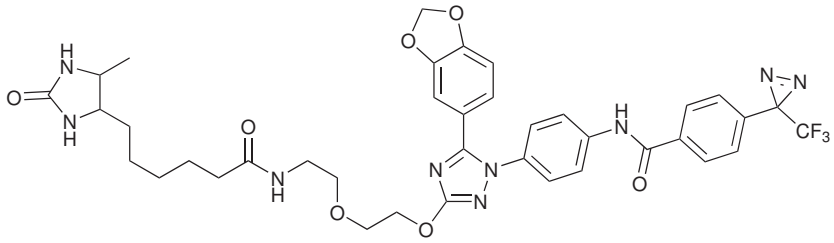

Figure 7.6: SecinH3-TPD - Chemical formula: $C_{38}H_{40}F_3N_9O_7$. Molecular weight: 791.8

4-(1-Azi-2,2,2-trifluoroethyl)benzoic acid (25 mg, 0.1 mmol, 1 eq.) was dissolved in THF (300 µl), and 2 drops of DMF added. The reactor was flushed with argon and a solution of oxalylchloride (92 µl, 1.0 mmol, 10 eq.) in THF (250 µl) added dropwise. The mixture was stirred for 2.5 h at room temperature and the solvent evaporated under vacuum. The oxalylchloride was removed by coevaporation with once THF and twice toluene. The raw product was dissolved in CH_2Cl_2 (500 µl) and the reactor purged with argon. **BA99** (63.2 mg, 0.1 mmol, 1 eq.) and *N*-ethyldiisopropylamine (38 µl, 0.2 mmol, 2 eq.) in CH_2Cl_2 (1.3 ml) were added dropwise and the reaction stirred for 2 h. The reaction mixture was concentrated under reduced pressure. Purification on TLC plates (MeOH/CH_2Cl_2 10:90) gave the desired product (35 mg, 41 %). Rf (MeOH/CH2Cl2/triethylamine 10:89:1) 0.66.

7.1.1.2 Solid-phase peptide synthesis

To get used to the LC-MS system a reference peptide was synthesised by solid-phase peptide synthesis on a PS3 peptide synthesiser (Protein technologies, Inc.). A 27 amino acids long subsequence of ARNO-Sec7 with two tryptic cleavage sites was chosen ("‖" denotes a tryptic cleavage site):

$$H_2N - FVAMNR \parallel GINEGGDLPEELLR \parallel NLYDSIR - CO_2H$$

Fmoc (fluorenylmethyloxycarbonyl chloride) protected amino acids and 0.4 mmol Rink Resin (PS3-RK-1) from Protein technologies, Inc. were used for the synthesis following the manufacturer instructions.

Deprotection

Various deprotection conditions were tested on aliquots. Then, the complete synthesis was deprotected following this protocol:

1. Prepare the deprotection solution: 82.5 % TFA, 2.5 % DTT, 5 % water, 5 % phenol, 5 % thioanisol.
2. Add 5 ml deprotection solution to 1 ml resin. Incubate 2 h at RT.
3. Filtrate over silica cotton.
4. Precipitate with diethylether.
5. Wash with diethylether.
6. Add 5 ml deprotection solution. Incubate 2 h at RT.
7. Precipitate with diethylether.
8. Wash 3-times with diethylether.
9. Dry on SpeedVac.

The peptide was then purified on a semipreparative HPLC. Column: Multo high bio 300–5 C4. Gradient: 30 % (FA/H_2O 0.1 % v/v)/CH_3CN to 100 % CH_3CN in ∼30 min. Injections: 50 µl, 1 mM. The collected fractions were pooled, shock frozen and dried on SpeedVac.

The purified peptide was analysed by LC-MS on the same column and with the following gradient:

min	FA/H_2O [%]	CH_3CN [%]
0	100	0
20	0	100
24	0	100
26	100	0

The peptide eluted at 15.5 min. Observed peaks: 773 and 1031 m/z.

7.1.2 Handling of nucleic acids

7.1.2.1 Agarose gel electrophoresis

Agarose gel electrophoresis was used to analyse the outcome of PCR reactions and enzymatic digestions. Concentrations of agarose varying from 0.5 to 2 % were used, depending on the expected DNA size. The desired amount of agarose was dissolved in 25 ml 0.5 X TBE-buffer by heating in a microwave oven and poured into the gel chamber. In order to detect the oligonucleotides bands, ethidium bromide at a final concentration of 0.5 µg/ml was added to the solution. After solidification, 5 µl of a 5:1 mixture of sample and 6 X agarose gel loading buffer were loaded and the gel run at 150 V in 0.5 X TBE-buffer. The oligonucleotide bands were then visualized on a transilluminator.

7.1.2.2 Determination of the concentration

The concentration of double stranded DNA solutions was determined from the absorption at 260 nm using the conversion rule: 1 $OD_{260} \equiv 50$ µg/ml DNA.

7.1.2.3 Molecular cloning

For cloning of new plasmid constructs the coding sequence of the protein of interest was amplified by PCR from a template plasmid. Both the amplified insert and a vector plasmid were digested with appropriate enzymes, mixed and ligated. The ligation reaction was used to transform ultracompetent cells which were cultured on antibiotic-containing plates. Then, colonies were picked an screened by PCR to verify the presence of the plasmid of interest.

Amplification of the insert

The inserts were amplified by PCR in 100 µl reactions:

Pipetting schema	
Water	to 100 µl
10 X Pfu rection buffer	10 µl
25 mM $MgCl_2$	8 µl
plasmid	2-10 ng
25 µM 5'-primer	2 µl
25 µM 3'-primer	2 µl
dNTP-Mix, 25 mM each	1 µl
2.5 U/ µl Pfu polymerase	4 µl

Cycling parameters		
95°	60 s	
95°	30 s	⎫
60°	30 s	28 X
72°	90 s	⎭
72°	180 s	
4°		

The PCR products were analysed by agarose gel electrophoresis, then purified with a PCR purification kit (Macherey-Nagel) and eluted in 30 µl 10 mM Tris, pH 8.0. The concentration was determined photometrically.

Enzymatic digestion of insert and vector

	Vector	Insert
water	to 20 µl	to 40 µl
10 X Fast Digest buffer	2 µl	4 µl
Vector	1 µg	-
PCR-product	-	850 ng
1 FDU/µl enzyme 1	0.5	1.5
1 FDU/µl enzyme 2	0.5	1.5

The reactions were incubated for 45 min at 37°C. Then, 0.2 µl CIP (Calf intestinal alkaline phosphatase, Promega) were added and the reaction incubated for further 15 min before purification with a MinElute Reaction Cleanup Kit (Qiagen).

Ligation

The digested samples were mixed in ratio of 3 mol insert to 1 mol vector. The whole insert sample was used. The mixture was incubated for 5 min at 45°C and then transferred directly on ice. 1/10 volume of 10 X ligation buffer (Fermentas) and 1 µl T4 DNA Ligase (Fermentas) were added and the reaction was incubated for 3 h at RT. The ligation reaction was used to transform XL10-Gold ultracompetent cells (Stratagene).

Transformation of XL10-Gold ultracompetent cells

Procedure:
1. Gently thaw the XL10-Gold ultracompetent cells on ice.
2. Add 1.25 µl of β-mercaptoethanol to 25 µl of the supercompetent cells.
3. Incubate on ice for 10 min (swirl gently every 2 min).
4. Transform with 10 µl of the ligation reaction.
5. Incubate for 30 min on ice.
6. Heat-pulse for 60 at 42°C.
7. Incubate on ice for 2 min.
8. Add 500 µl of preheated SOC medium to each tube and incubate for ∼60 min (depending on the antibiotic) at 37°C with shaking (Thermomixer, 2 ml, 1000 rpm)
9. Centrifuge the cells for 3 min, 3000 rpm, discard ∼450 µl of the supernantant.
10. Plate the cells (∼75 µl) on agar plates with the right antibiotic and incubate o.n. at 37°C.

Usually, 10 colonies were picked, boiled for 5 min in 25 µl water and centrifuged at 15 000 rpm for 3 min. 5 µl of the supernatant were mixed to 5 µl of master mix and used for colony PCR. A primer was chosen to be in the vector and the other in the insert. The clones whih resulted positive in the colony PCR, were inoculated in 5 ml LB medium with appropriate antibiotic and grown o.n. The plasmid were minipreparated and the coding sequence verified by sequencing (GATC Biotech).

Master mix for 11 PCRs	
Water	42.5 µl
10 X Taq buffer	12 µl
dNTP mix, 25 mM each	0.96
25 µM 5'primer	0.96 µl
25 µM 3'primer	0.96 µl
5 U/µl Taq	2.4 µl

Cycling parameters		
95°C	60 s	
94°C	30 s	⎫
60°C	30 s	25 X
72°C	60 s	⎭

7.1.2.4 Mutagenesis

The QuikChange lightning site-directed mutagenesis kit was used to obtain plasmids encoding the GEF inactive ARNO mutant E156K.

Plasmids used (10 ng/µl): ARNO GFP K7 (Resistance: zeocin)
pCMV-FLAG Cyt2 Full (Resistance: kanamycin)
pCMV-FLAG Cyt2 Sec7 (Resistance: kanamycin)
5'-Primer: ARNO-E156K sp, 15 µM
3'-Primer: ARNO-E156K ap, 15 µM

Pipetting schema:	
water	34.5 µl
10 X reaction buffer	5 µl
Plasmid	5 µl
5'-Primer	1 µl
3'-Primer	1 µl
dNTP mix	1 µl
QuickSolution	1.5 µl
Lightning Enzyme	1 µl

Cycling parameters:		
95°C	120 s	
95°C	30 s	⎫
60°C	60 s	18 X
68°C	165 s	⎭
4°C		

2 µl DpnI were added and the reaction incubated at 37°C for 30 min before purification with a PCR purification kit (Qiagen). The whole reaction was used to transform XL10-Gold ultracompetent cells (Stratagene).

7.1.3 Protein methods

7.1.3.1 Determination of protein concentration

The concentration of purified protein solutions was measured by UV absorption at 280 nm and calculated according to the Beer-Lambert law

$$A = \epsilon \, \ell \, c$$

where c is the concentration in mol/l, A the absorbance, ℓ the path length of the sample in cm. The molar extinction coefficient ϵ at 280 nm is given in $l\,mol^{-1}\,cm^{-1}$ and calculated with VectorNTI, based on the content of tyrosine, tryptophan, and cysteine residues.

Bradford assay

The protein concentration in complex samples (e.g. cell lysate) was determined by Bradford assay[152]. This colorimetric assay bases on the binding of the Coomassie Brilliant Blue dye to basic and aromatic amino acids which leads to an absorbance shift of the absorption maximum to $\lambda = 595\,\text{nm}$. The assay was measured in 96-well microtiter plates. BSA dilutions (333, 500, 666, 1000, 1500, 2000 and 3000 µg/ml) were used to build a calibration curve. Each 2 µl of the sample of interest and the BSA dilutions were dispensed in the wells (in triplicates and duplicates, respectively) and then 150 µl of 1:6 diluted Bradford solution (Bio-rad) were added. The absorption at 595 nm was measured in a Varioskan plate reader and the samples concentration derived from the calibration curve.

7.1.3.2 SDS-PAGE

For the analysis of protein purity, expression or labelling, the protein samples were separated according to their size by discontinuous SDS-PAGE (sodium dodecyl sulfate polyacrylamid gel electrophoreses) according to Laemmli[153].

The gels were prepared by polymerisation of acrylamide and *N,N'*-methylene-bis-acrylamide initiated with ammonium persulfate (APS) and *N,N,N',N'*-tetramethyl-ethylenediamine (TEMED) according to the following schema:

	Separating gel		**Stacking gel**
	7.5 %	**10 %**	**4 %**
(Bis)-acrylamide	1250 µl	1667 µl	213 µl
Water	2462 µl	2045 µl	975 µl
Separating gel buffer (4x)	1250 µl	1250 µl	-
Stacking gel buffer (4x)	-	-	400 µl
TEMED	8 µl	8 µl	2 µl
APS	30 µl	30 µl	10.4 µl

The indicated reactions were mixed and the gels were cast between two thin glass slides. The concentration of the separating gel was chosen depending on the size of the investigated protein. Usually 7.5 % gels were used for the analysis of ARNOΔPBR and 10 % gels for the analysis of ARNO-Sec7 and the EGFR. A stacking gel, poured on top of the separating gel, was used to focus the proteins in a thin starting zone. The samples were boiled in 1 X loading buffer (p. 148) loaded on the gel and separated for ∼1 h at 200 V in cold Tris glycine running buffer (p. 147). Page Ruler Prestained Protein Marker (Fermentas) was loaded as a size standard. The gels where then either stained with Coomassie or transferred on a nitrocellulose membrane by Western blotting.

7.1.3.3 Western blotting and immunodetection

Discontinuous blotting according to Kyhse-Andersen[154] was used to transfer proteins separated by SDS-PAGE on nitrocellulose membranes. The gel was piled between

different layers of filter papers equilibrated for some minutes in different buffers (p. 148) as following:

Cathode
3 filter papers in cathode buffer
Gel in Cathode buffer
Nitrocellulose membrane in anode II buffer
2 filter paper in anode II buffer
1 filter paper in anode I buffer
Anode

A constant power of 2 mA/cm^2 gel was applied for 45 min. Then, the membrane was blocked by incubation in 5 % w/v milk powder or BSA in 1 X TBS-T for at least 45 min at RT, rinsed 3-times with 1 X TBS-T and incubated with dilutions of primary antibody either for 1 h at RT or o.n. at 4°C. The membranes were washed 3-times in 1 X TBS-T and then incubated in horseradish peroxidase (HRP)- or near infrared (NIR) dye-coupled secondary antibody diluted 1:10 000-1:20 000 in 5 % w/v BSA in 1 X TBS-T for 1 h at RT. NIR dye coupled antibodies were incubated in the dark.

Visualisation was done by enhanced chemiluminescence on a VersaDoc 5000 CCD camera (BioRad) or by NIR immunofluorescence on an Odyssey scanner (Licor). The bands were quantified with the QuantityOne software (BioRad).

7.1.4 Cell culture

Cells were grown at 37°C and 5 % CO_2 in the following media and splitted every 2-3 days, when they reached a maximal confluency of ∼80 %.

HEK293T	DSMZ Cell Lines Bank	DMEM, High Glucose, with phenol red and 10 % FBS	PAA/Lonza
H460	ATCC Cell Biology Collection	RPMI (PAA) with 10% FBS	PAA/Lonza

7.1.4.1 Preparation of cell lysates

For the analysis of EGF stimulation, the transfected HEK cells (in 6-wells plates) were serum-starved overnight in DMEM, 0.1 % FBS and stimulated for 5 min with 50 ng/ml EGF (Peprotech). The medium was discarded and the cells harvested in cold 1 X PBS, 5 mM EDTA on ice. After centrifugation (500 rpm, 5 min, 4°C) the supernatant was decanted, the pellets resuspended in ∼30 µl lysis buffer supplemented with protease-inhibitor-mix HP (Serva) and incubated for 15 min on ice. Cell debris were separated by centrifugation at 12 000 rpm for 20 min at 4°C) and the protein content of the supernatant determined by Bradford assay. The samples were diluted in water and 6 X Loading buffer to ∼3 µg/µl in 1 X Loading buffer and boiled. 40-50 µg proteins were loaded for SDS-PAGE.

1 X Lysis buffer	20 mM	Tris, pH 7.5
	150 mM	NaCl
	1 mM	EDTA
	1 mM	EGTA
	2.5 mM	sodium pyrophosphate
	1 mM	β-glycerophosphate
	1 mM	sodium vanadate
	1 % (v/v)	Triton X-100

7.1.5 BRET

For BRET, HEK cells transiently expressing the fluorophore/enzyme-tagged proteins were used. The plasmids transfected for each kind of experiment were:

Experiment	Donor	Acceptor	Ratio	Additional plasmid
Direct	pEGFR-RLucd4	pARNO-GFP2	1:20	
Indirect	pEGFR-RLucd4	pEGFR-GFP2d4	1:20	pCMVTag2-ARNO
	pHER2-Luc	pHER3-GFP	1:20	pCMVTag2-ARNO
Control	pGST-Rluc	pGST-GFP2	1:20	pCMVTag2-ARNO

7.1.5.1 Transfection

3.5×10^6 HEK293T cells were seeded in a 10-cm plate in 10 ml DMEM with 10 % FBS and without phenol red. Metafectene (Biontex) (3 µl per µg of DNA was premixed with serum free media. Maximally 6 µg total DNA (transfections with less DNA were filled up with empty vector) were diluted in serum free medium and then added to the metafectene solution without mixing. After 20-30 min the cells were transfected by addition of the transfection mixture. The cells were then incubated for ∼40 h at 37°C and 5 % CO_2.

7.1.5.2 Measurement

Solutions needed:

BRET-buffer: add 0.55 g glucose to 500 ml DPBS (PAA), warm up to 37°C

Luciferase substrate: 15 µM DeepBlueC in PBS. Dilute from the stock solution (1 mM in ethanol) only at the moment of use*. 50 µl/well + 600 µl for priming are needed.

5 mM EDTA in PBS to detach the cells.

Before start:

warm up the BRET-buffer and a white 96-wells plate to 37°C.

Start the instrument to allow warming up, open the parameters file to set the temperature to 37°.

Wash the injector with water ('Instrument – Wash', tick 'injector 3', press

* DeepBlueC decays quickly. Keep the time between dilution and injection as short as possible.

'next' twice).

1. Take the cells from the incubator and remove the medium.
2. Detach the cells with 1 ml DPBS/EDTA and transfer into a 15 ml-Falcon.
3. Wash the plate with 6 ml BRET-Buffer and add to the Falcon. Take an aliquot to count the cells.
4. Centrifuge at 400 rcf, 37°C, for 5-7 min. Meanwhile, count the cells.
5. Decant the supernatant.
6. Add BRET-Buffer to a final concentration of 3.2×10^6 cells/ml.
7. Mix and transfer in 8-ates (in columns) into a white 96-well plate (100 µl/well, keep warm)
8. Dilute the luciferase substrate. Connect to the instrument.
9. Prime the instrument ('Instrument – Prime', tick 'injector 3', press 'next' twice).
10. Select 'Platte einstellen' (bottom left), tick 'by rows', select the wells.
11. Write a filename.
12. Press Start to insert the plate.
13. After the measurement: 'Datei – Export', 'Matrix export'. The report is saved in 'Dokumente – Mithras Datei – Transfer'.
14. Give two aliquots of the cells in a black plate and measure total GFP.
15. 'Instrument – unload plate'.
16. Substitute the luciferase substrate with water. 'Instrument – wash', select 40 cycles, 'next'.

7.1.5.3 Evaluation

1. Check that luciferase signal is over 1000.
2. Subtract background from non trasfected cells.
3. Calculate ratio GFP/Luc
4. Subtract ratio of cells transfected with Luciferase only.

7.1.6 SPR

7.1.6.1 Maintenance routines

SPR machines are particularly sensitive to impurities and especially proteins tend to adsorb to the tubing. To ensure optimal performance of the machine, the following working rules were defined:

- Buffers are always filtered and degassed and possibly stored at 4°C. Check regularly that nothing is growing in the buffer. Consider making aliquots of your buffers to reduce contaminations.

- Execute a DESORB routine at least once a month and before every chip coupling. If a new interaction is studied, it is worth running a DESORB routine before start.

> **!** Never run a DESORB, SANITIZE or SUPERCLEAN procedure on your working chip as this could be destroyed. Use the maintenance chip instead.

- Run a SANITZE routine at least every three months. Always execute a DESORB prior a SANITIZE.

  > **!** After SANITIZE a longer rinsing time is needed. It is worth planning the routine at the end of the day.

- From time to time run a SUPERCLEAN by priming the system four times in a row with each of the following reagents: 0.5 % SDS, 6 M urea, 1 % acetic acid and 0.2 M $NaHCO_3$.
- Run on occasion a SYSTEMCHECK to verify the condition of the IFC.
- Prime the system three times at every buffer change (it is time well invested!).
- Let the chips ad buffers warm up to RT for ∼30 min before use.
- Store your ligand coupled chips wet (especially if protein is coupled). I.e. after undocking the chip remove the chip carrier slide from its sheath and insert it in a 50 ml Falcon tube filled with your running buffer. Store at 4°C.
- Clean your chip before redocking. I.e. wash the plastic slide and the gold surface with water and dry it with compressed air. Make sure that no salts deposit are visible on the gold surface.
- If possible, test the solubility of the molecules before injecting them. If your analyte aggregate/precipitate in the system, you can spend the rest of the day with cleaning procedures!

7.1.6.2 Analysis of protein/inhibitor interactions

The SPR measurements of the ARNO/inhibitor interaction were performed by Esteban Gutierrez during his Master thesis[76]. An adaptation of his methods description is reported here for completeness.

Preconditioning of SPR chips

The CM5 chip for the second coupling (Fig. 4.24) was preconditioned before covalent coupling of the small molecules. The chip was docked in and the system primed with water. A new sensorgram was started, with a flow rate of 100 µl/min. 100 mM HCl, 50 mM NaOH, 0.5 % SDS w/v and water were injected sequentially for 10 s each. The sequence was repeated once. Water injections where repeated until a stable baseline was observed.

Coupling of small molecules

For the coupling of small molecules on CM5 sensor chips standard EDC/NHS coupling was used following the manufacturer instructions. The compounds were derivatized with amine linkers as described in Section 4.3.2.3 (p. 55). A sodium hydrogenocarbonate buffer (pK_a 10.3) was used to ensure that the primary amines (pK_a 9-10) were not protonated. 5 % DMSO were added to avoid solubility problems. As running buffer HBS-P (Biacore) was used.

Coupling protocol:
1. Allow the CM5 chip and all buffers to warm up to RT.
2. Prepare fresh EDC (0.4 M in water) and NHS (0.1 M in water) solutions or thaw up frozen aliquots.
3. Dock the chip and prime with HBS-P.
4. Start a new sensorgram on Fc2 (Fc1 is used as reference cell). Flow rate: 5 µl/min .
5. Activate the surface by a 7 min injection of a 1:1 mixture of the EDC/NHS solutions.
6. Immobilize the first small molecule: 10 min injection (150 µM in 10 mM Na_2CO_3, pH 10.5, 5 % DMSO).
7. Deactivate the surface with a 7 min injection of ethanolamine solution. Stop the sensorgram.
8. Repeat steps 4.-7. for Fc2 and Fc3 with the other small molecules. Stop the sensorgram.
9. Start a new sensorgram on Fc1 (reference cell).
10. Activate the surface with a 7 min injection of EDC/NHS 1:1.
11. Inject buffer (10 mM Na_2CO_3 pH 10.5, 5 %DMSO) for 5 min.
12. Deactivate with a 7 min injection of ethanolamine solution. Stop the sensorgram.

With this protocol 2 000-3 000 RU of small molecule where coupled on the untreated CM5 chip and around 800 RU on the preconditioned chip.

Acquisition of binding sensorgrams

Before every interaction kinetic measurements, after docking the ligand-coupled chip and priming the system with running buffer, the UNCLOG and RINSE routines (Biacore control software - working tools) were used. This reduces artifacts from air bubbles and contaminations of IFC, injection needle or chip.

To monitor the surface and injection stability, a sensorgram was started on the desired flow cells (reference and ligand-bound Fcs) and run with a flow rate of 50-100 µl/min for 5-15 minutes. The slope of the baseline was determined by setting report point. Baselines with slopes lower than 0.02 RU/s were considered as stable. If stronger drifting was observed, the systems working tools were used (CLEAN, RINSE,

UNCLOG) and, if necessary, regeneration solutions injected (5 mM NaOH, 1 M NaCl, 0.5 % SDS or 5 mM glycine, depending on the sensitivity of the ligand).

We frequently observed irregularities in the first few injections. Therefore, after observing baseline stability, we executed several buffer injections (INJECT or KINJECT mode, 60 s injections) until a clean injection sensorgram (square shaped or flat) was observed.

The kinetic measurements were performed in the KINJECT mode, at a flow rate of 30 µl/min with ARNO buffer (p. 148) as running buffer. The KINJECT injection routine avoids pumps filling during the dissociation phase (pump filling often lead to sensorgram jumps which disturb analysis). 250 µl of protein solution were injected and then the dissociation monitored for 5 min. For double referencing[66], two buffer injections with the same parameters were acquired before protein injection on one again at the end of the measurement series. Two injections of buffer were also done (INJECT, 60 seconds) before each injection of analyte sample, to avoid irregularities at the start of the sensorgram (due to residual regeneration solution in the needle).

Five concentrations (50 nM, 100 M, 250 nM, 500 nM, 1 µM , 1.5 µM) of protein solution (ARNO-Sec7 and GST) were injected. Every sample injection was followed by 3 regeneration injections (5 mM NaOH and 500 mM NaCl; 0.5 % SDS; 5 mM glycine), and a new cycle was started between the regeneration phase and the following analyte injection.

Protocol used:
1. Change running buffer to ARNO buffer.
2. Start a new sensorgram: mode "FC1,2,3,4", flow rate 30 µl/min. Wait until a stable baseline is reached.
3. Set data collection rate to high.
4. INJECT running buffer to test surface stability.
5. Acquires two KINJECT with buffer for double referencing.
6. Start a new cycle. INJECT buffer for twice 60 s
7. Use KINJECT for injections. Set the association time to 250 µl and the dissociation time to 5 min.
8. KINJECT the first concentration of protein.
9. INJECT the regeneration solutions. Try different regeneration conditions until the baseline is back at the level measured before analyte injection.
10 Repeat steps 6.-9. for the remaining concentrations.
11. Acquire a KINJECT with buffer for double referencing.

Data processing and analysis

Sensograms obtained from kinetic interaction measurements were double referenced[66]. An example of data processing is shown in Figure 7.7.

First, the response in the few seconds just before injection was set to zero (y-axis) and the start of the injection was set to zero by eye for each each Fc (Fig. 7.7**b**).

Then, the reference surface data is subtracted from the reaction surface data to remove the refractive index change and injection noise (Fig. 7.7**c**). The sharp spikes at either the beginning or end of the association phase are due to differences in the position of the reaction and reference Fcs and should be removed during data analysis.

The reference subtracted response and the blank injections (buffer KINJECTions) are overlaid (Fig. 7.7**d**) and the response from the blank injections subtracted (Fig. 7.7**e**) and, at the end, the responses from the different analyte concentrations are overlaied (Fig. 7.7**f**).

At this point the dissociation is determined with the program BIAevaluation (Biacore) by global kinetic curve fitting to a 1:1 interaction model or equilibrium analysis.

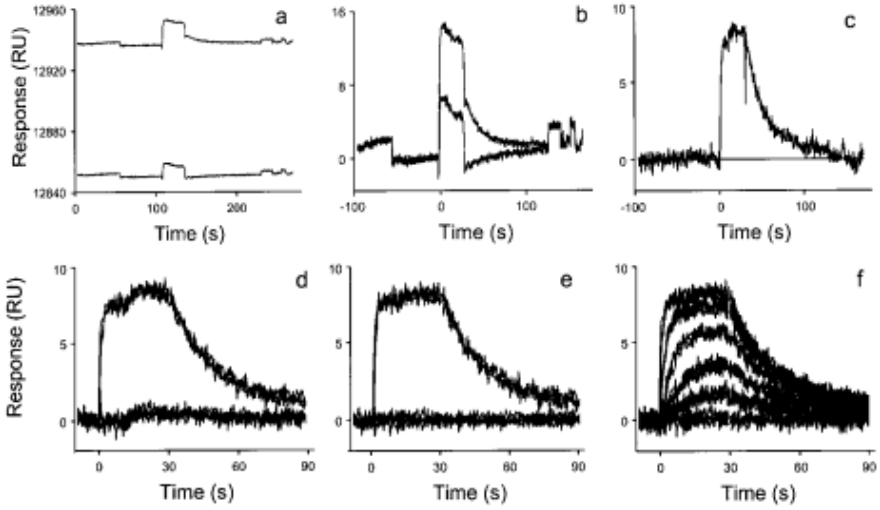

Figure 7.7: Example of data processing - **a.** Raw data from biosensor for 233 nM IL-2 injected over a receptor surface (top trace) and reference surface (bottom trace). **b.** Data sets were zeroed on the y and x axis just prior to the start of the injection. **c.** Data from the reference surface was subtracted from the data from the reaction surface. **d.** Overlay of four replicate injections of 233 nM IL-2 as well as a running buffer blank. **e.** After subtracting the average of the blank injections from both the sample and blank data. **f.** Overlay of a series of IL-2 injections (233, 78, 26, 8.6, 2.9 and 0 nM) replicated four times each.[66] *Adapted from Journal of Molecular Recognition (Ref. 66), copyright 1999, with permission from John Wiley and Sons, Inc.*

7.1.6.3 Analysis of protein/protein interactions

Loading of the NTA/NiP chips

NTA chips were first washed with 0.5 M EDTA, pH 8.4, to get rid of Ni^{2+} ions and then loaded with 20 mM $NiCl_2$. To avoid carry-over of nickel, 3 mM EDTA was added to the dispenser buffer (p. 148) (connected to pump B and used for sample preparation

and washing of the injection needle) and another rinsing programs 'EXTRACLEAN' or 'WASH Needle' were used (the routine 'WASH IFC' has increased risk of carryover).

The His-tagged protein was injected and NTA running buffer flown, till a stable signal was observed. Then, the analyte was KINJECTected at 30 µl/min . The KINJECT injection routine avoids pumps filling during the dissociation phase (pump filling often leads to sensorgram jumps which disturb analysis).

The chip was regenerated by repeated injections of NTA regeneration buffer (p. 148) and 0.5 % SDS in 10 mM HEPES, pH 7.0, 150 mM NaCl and 0.005 % Tween-20. High salt buffers were used to reduce electrostatic interaction of the proteins with the chip surface.

Covalent immobilisation on NTA chips

The NTA chip was first washed with 0.5 M EDTA, pH 8.4, to get rid of Ni^{2+} ions. The cell of interest was loaded with 20 mM $NiCl_2$ (2 min) and activated with a fresh 1:1 mixture of 0.4 M EDC in water and 0.1 M NHS in water (3 min). Then, 100 nM His-EGFR in NTA running buffer were injected (10 min) and the surface was deactivated by a 3 min injection of 1 M ethanolamine hydrochloride/NaOH pH 8.5. The Ni^{2+} ions and the non covalently bound receptor were washed away with NTA regeneration buffer (3 min).

The analyte was KINJECTected in different running buffer at 30 µl/min as described in Section 4.3.1.3 (p. 44). Various regeneration solutions were used as described in the previous section.

Covalent coupling to CM5 chips without preconcentration

The CM5 chip was activated with a fresh 1:1 mixture of 0.4 M EDC in water and 0.1 M NHS in water (6 min). Then, 100 nM GST-EGFR in NTA running buffer were injected (15 min) and the surface was deactivated by a 2 min injection of 1 M ethanolamine hydrochloride/NaOH pH 8.5. The analyte was KINJECTected at 30 µl/min in HBS-P. The most promising regeneration solution was 0.1 % SDS/1 % Triton X-100 in HBS-P, but still complete regeneration was not possible.

7.1.7 MS and LC-MS analysis

7.1.7.1 Maintenance routines

To ensure optimal performance of the machine the following should be respected:

- Leave the system running with nitrogen: ≥ 5 l/min of dry gas flow. Switch the instrument to 'Shutdown' for maintenance only. When not in use keep in 'Standby'.
- Control regularly the nitrogen and helium pressure.
- Check detector calibration at least one a month.

- Note in the logbook:
  - the dates of detector calibration and current multiplier voltage.
  - the vacuum values at least once a month and before and after system venting.
  - the Accu Time (right click in the MS spectrum on the right, select 'Eigenschaften' and tick 'Show Accu Time') and capillary current (visible in the 'Tune' tab under 'Capillary') at every detector check.
  - the dates of maintenance and other actions.
- Measure regularly a standard sample with a standard LC-MS method to test the whole set up.
- Clean the spray chamber and spray shield weekly.
- Use acetonitrile (ACN) of at least HPLC grade and water of MilliQ quality.
- Avoid tubes containing plasticiser. Do not wash the bottles in the dish washer, rinse them with clean solvents instead.

Cleaning

Clean the spray chamber and spray shield weekly with isopropanol/water.

If the accumulation times in MS mode increase (check during calibration) and the signal intensities drop, the system is probably contaminated and needs cleaning. A longer run with alternating polarity can help cleaning the system:

> Wash o.n. with 200 µl/min 80 % ACN/H_2O. Untick 'HV', set the 'Dry gas' to 10 l/min and the 'Nebulizer' to 30 psi. Thick 'Alternating' under 'Polarity'.

If the signal does not improve, manual cleaning is needed. Refer to the manual.

Adjusting the spray needle

For best ESI performance and MS sensitivity the tip of the nebulizer should be checked from time to time for correct position, damage and contamination:

1. Switch the 'Neb Gas' off.
2. Detach the gas inlet and open the spray chamber.
3. Take out the nebulizer and place it in the support block.
4. Position the magnifier so that you can see the nebulizer tip.
5. If necessary, loosen the needle holder locknut.
6. Adjust the needle until it sticks out about 1/2 of its own diameter.
7. Tighten the locknut while making sure that the position of the needle does not change.

! Be careful with the tip of the nebulizer needle. It is very sensitive and can be damaged when touched.

If the needle need to be cleaned or replaced, please refer to the manual.

Calibration

The detector calibration was checked weekly and recalibration performed when necessary:

1. Dilute 10 µl 'ES tuning mix' (Agilent) with 190 µl acetonitrile and 10 µl water.
2. Inject by infusion at 240 µl/min.
3. Set the default method for MS acquisition (press Ctrl+N in the esquire-Control window to open the default method).
4. Wait till a stable signal is observed. If the signal does not stabilise flush with 50 acetonitrile/water at higher flow rate (by hand).
5. In the 'Calibration' tab, press 'Detector...'. Select 'ES Tuning Mix Pos' as 'Mass list' and 1521.97 m/z as 'Mass'.
6. Press check and wait till the results windows pops up
7. If the program ask for recalibration, press 'Calibrate' in the 'Detector' window.
8. 'Check' again after calibration.
9. Note the dates of detector calibration, current multiplier voltage, Accu Time, capillary current and intensity of the signal at 1522 m/z in the logbook.

Calibration of the masses is usually not necessary. If you notice a mass shift when measuring the standard sample, perform a Scan calibration with the same calibration solution:

1. Follow steps 1.-4. of the detector calibration protocol.
2. In the 'Calibration' tab, select 'Auto' and then press 'Start'. Follow the instructions.

Standard sample

The whole set up should be checked regularly with a standard sample. I used 5 pmol of trypsin-digested BSA separated on C18-column with the LCMS-BSA-2-40-30min (p. 138) method.

The digested BSA was either bought from BioLabs or prepared by digestion in solution (p. 143) of BSA.

7.1.7.2 LC-MS programs

For LC-MS analysis the following 125 × 4 mm columns from CS-Chromatographie Service GmbH were used:

MultoHigh-Bio 300 – C4 5µ, Art.-Nr. 584 1177, Column-Nr. 3303-10 (Analysis of non digested proteins)
MultoHigh 100 RP 18 – 5µ, Art.-Nr. 582 1172, Column-Nr. 12911-05 (Analysis of digested proteins)

The following parameters were used for MS acquisition:

Scan modus	Standard-Enhanced, positive
Neb	30 psi
Dry gas	10 l/min
Dry temp	330°C
Smart Target	200 000
Scan	300 to 1500 m/z
Averages	5
Target mass	900
Auto MS(2) acquisition parameters:	
Scan	100 to 2300 m/z
Averages	3
Frag Ampl	0.5 V
Active exclusion:	Excluded after 3 spectra
	Release after 0.5 min

LC gradient for protein analysis

Undigested proteins were diluted in 30 % isopropanol in 0.1 % FA/H_2O and analysed with the following gradient on the C4 column:

min	**0.1 % FA/H_2O** [%]	**Isopropanol** [%]
0	70	30
3	70	30
23	40	60
24	0	100
29	0	100
30	70	30

MS spectra were acquired without MSMS between 5 and 25 min from gradient start.

Peptide mixtures analysis

Digested samples of purified proteins were usually analysed with the following 30 min gradient on the C18 column:

min	0.1 % FA/H_2O	ACN
0	98	2
30	60	40
31	10	90
36	10	90
37	98	2
60	98	2

MS or MSMS spectra were acquired between 3 and 18 min from gradient start.

Instead, digested samples from the enrichment experiments were analysed with a 60 min gradient on the C18 column:

min	0.1 % FA/H_2O	ACN
0	98	2
3	98	2
40.2	53	47
42	0	100
47	0	100
49	98	2
70	88	2

MS or MSMS spectra were acquired between 8 and 40 min from gradient start.

7.1.8 CCMS

7.1.8.1 Labelling reactions

Alkylation with iodoacetamide

Alkylation with iodoacetamide (IAA) routinely used in protein digestion protocols to avoid reformation of disulfide bonds after denaturing and reduction. In this case we were interested in introducing a modification in ARNO which could be detected by MS. ARNO-Sec7 has 3 cysteines, which are all found on the same tryptic peptide.

The protocols for alkylation found in literature use highly different incubation times and temperatures. Since many side reactions are possible[155], I decided to use a short alkylation time at room temperature. The protocol was adapted from Ref. 91:

Solutions needed: 12 M urea in water. Warm up to 50 °C to dissolve.
1 M IAA in water.

1. Dilute ARNO-Sec7 to 10 µM in 50 mM Tris pH 8.5 (total volume 50 µl).
2. Add 50 µl 12 M urea and 1 µl 1 M DTT.
3. Incubate 25 min at 65 °C.
4. Add 4.2 µl fresh 1 M IAA.
5. Incubate 15 min at 25 °C in the dark.
6. Buffer exchange with centrifugal filters (Amicon Ultracell 10 K - 0.5 ml) (see next section for an example).

7. Dilute to 5 µM with 10 mM NH_4HCO_3.
8. Add 0.4 µl trypsine (1 µg/µl) and incubate for 13 h at 37°C.
9. Dilute to 1 µM with 0.1 % formic acid/water.

For LC-MS analysis of the undigested protein, an aliquot from step. 6. was diluted to 0.1 µM with 0.1 % formic acid/water and 10 % Isopropanol were added. 35 µl were injected on a C4 column as described in Section 7.1.7.2. The digested sample was analysed by injection of 60 µl of the step 9. dilution on a C18 column with the 60 min programm (p. 138).

Biotinylation with PEG2-biotin

EZ-link iodoacetyl-PEG_2-biotin (IAA-biotin) was used to biotinylate ARNO which was used as reference in western blot and for the optimisation of the enrichment procedure. An example of an ARNO-Sec7 biotinylation protocol is given:

1. Dilute ARNO-Sec7 to 10 µM in 50 mM Tris pH 8.5 (total volume 65 µl).
2. Add 65 µl 12 M urea and 1 µl 1 M DTT.
3. Incubate 25 min at 65 °C.
4. Add crumb* of IAA-biotin.
5. Incubate 20 min at 25 °C in the dark.
6. Buffer exchange with centrifugal filter (Amicon Ultracell 10 K - 0.5 ml) to labelling buffer without glycerol:
   Fill up the sample to 500 µl and transfer to the filter.
   Centrifuge at 11 000 g, 5 min.
   Add ∼400 µl buffer, centrifuge at 11 000 g, 5 min. Repeat 3 times.
   Invert the filter in a new tube. Elute by centrifuging at 1020 g, 3 min.

SecinH3-TPD labelling of purified protein

SecinH3-TPD was conserved dried at 4°C in the dark. Aliquots reconstituted to 25 mM in DMSO were conserved at room temperature for several months. Further dilutions were done directly prior usage and discarded after the experiment.

! Avoid light exposition of SecinH3-TPD.

All the labelling reactions were performed on ice with prechilled labelling buffer (50 mM Tris pH 7.65, 300 mM NaCl, 3 mM $MgCl_2$, 10 % Glycerol) with or without 0.005 % Triton X-100 as indicated. The proteins were prediluted in labelling buffer and aliquoted in 1.5 ml test tubes. The compounds were prediluted in DMSO or Diglyme, added to the protein dilution and mixed by pipetting

! Pipetting order: to avoid compound precipitation it is important to add the compound dilution to the protein solution.

*The amount needed is far too low to be weighted. If you can see it, it is enough.

The samples were incubated in the dark for 10 min and then irradiated for 10 min with UV light at 365 nm (3UV Lamp, bulb power 8 W, UVA output at 1 cm 2800 µW/cm^2).

For western blot analysis aliquots were taken before and after irradiation, mixed with 1/6 volume of 6X loading buffer (p. 148) and boiled for 5 min. The samples were separated by PAGE and transferred on membrane by western blotting (p. 127). The membrane was blocked for 1 h with 5 % BSA/TBS-T w/v (p. 148).

! Do NOT use milk to block the membrane if biotin detection is following.

For detection of total protein, the membrane was incubated with His_5-antibody (diluted 1:2 000 in 5 % BSA/TBS-T w/v with 0.02 % thimerosal) for 1 h, washed 3-times 5 min with TBS-T and then incubated for 30-60 min with near infrared (NIR) dye or horseradish peroxidase (HRP) coupled secondary antibody. After washing 3-times 5 min with TBS-T, visualisation was done by enhanced chemiluminescence on a VersaDoc 5000 CCD camera or by NIR immunofluorescence on an Odyssey scanner.

For detection of labelled protein, the membrane was incubated with either NeutrAvidin DyLight 800 (1:40 000 in 5 % BSA/TBS-T w/v) for 30-60 min or HRP coupled streptavidine (1:10 000 in 5 % BSA/TBS-T w/v), washed at least 4-times 5 min with TBS-T and visualised as above. If the NeutrAvidin DyLight 800 treated blot showed high background an additional washing step of up to 2 h was done.

! Incubate NIR dyes dilutions in the dark.

Example of a standard labelling reaction

End concentrations: 1 µM ARNOΔPBR, 20 µM SecinH3-TPD, 10 % DMSO.

1. Dilute ARNO to 1.1 µM in labelling buffer with 0.005 % Triton X-100. Keep on ice.
2. Aliquot 49.5 µl in a new cold tube. Keep on ice.
3. Dilute SecinH3-TPD to 125 µM in DMSO at room temperature.
4. Add 5.5 µl of the dilution in 3. to the aliquot in 4. Mix well by pipetting. Incubate on ice for 10 min.
5. Transfer 25 µl in a new cap, add 5 µl 6X loading buffer. Boil 5 min.
6. Irradiate for 10 min at 365 nm on ice.
7. Transfer 25 µl in a new cap, add 5 µl 6X loading buffer. Boil 5 min.
8. Load each 12 µl on a 7.5 % polyacrylamide gel.

SecinH3-TPD labelling in proteome

To test the binding and labelling specificity of SecinH3-TPD, labelling reactions in were performed in cell lysate. Lysis of HEK or H460 cells followed this protocol:

Lysis buffer: 50 mM Tris pH 7.65, 300 mM NaCl, 3 mM $MgCl_2$, 25 µg/ml digitonin (mix 40 µl 10X labelling buffer (p. 149), 11.1 µl digitonin (0.9 mg/ml in water) and 349 µl water).

1. Collect cells grown to confluency in a 150 cm^2-flask (often, frozen pellets were used).
2. Resuspend the cells in 200 µl lysis buffer.
3. Incubate 1 h on ice. Mix by vortexing from time to time.
4. Centrifuge 30 min at 4°C, maximal speed, to separate the cell debris.
5. Transfer the supernatant in a new tube and measure the protein concentration by Bradford assay (p. 127).
6. Use in the labelling reaction at a final concentration of 2 mg/ml. If the sample is needed for the enrichment experiment (next section), use 1 µg proteome/sample.

The labelling reaction was performed as described above. The samples for enrichment were aliquoted in 2 2-ml tubes (500 µl each) before irradiation and irradiated for 15 min instead of 10 min. Labelling buffer without glycerol and with 0.005 % Triton X-100 was used.

7.1.8.2 Enrichment of labelled proteins and digestion on beads

Biotinylated proteins were enriched by precipitation with streptavidine beads (Pierce Streptavidin UltraLink Resin) starting from labelled proteome (Section 7.1.8.1) or cell lysate with added prebiotinylated protein. Trypsin digestion was done on the beads and the supernatant collected for MS-analysis. The biotinylated peptides were eluted from the beads with acetonitrile/water. The protocol was adapted from Ref. 92.

! Do not use Dynabeads M-280 Streptavidin. Since they are conserved in a BSA containing buffer, high BSA signals are disturbing the MS analysis. If working with magnetic beads is preferred, BSA free beads should be used (e.g. Dynabeads MyOne Streptavidin C1, Invitrogen).

Solutions needed: PBS
1.2 % SDS/PBS
6 M urea/PBS (prepare fresh daily)
6 M Urea/PBS with 10 mM DTT (prepare fresh before use)
440 mM IAA in water (prepare fresh before use)
100 mM $CaCl_2$ in water

Enrichment

1. Pool the labelling reaction (1 ml) in a 15 ml Falcon tube.
2. Add 4 ml of cold (-80°C) acetone. Incubate at -80°C overnight (shorter time can be used).
3. Centrifuge at maximal speed at 0°C for 45 min.
4. Discard the supernatant. Let the pellet dry for ∼15 min at RT.
5. Dissolve in 1 ml 1.2 % SDS/PBS by repeated cycles of sonication and boiling.
6. Add PBS to 12 ml.

7. Wash 100 µl of beads slurry 3-times with PBS (collect the beads by centrifugation at 1400 g, 3 min **!**Always use the soft brake function when centrifuging the beads**!**).
8. Resuspend in 50 µl PBS and transfer in the Falcon.
9. Incubate 3 h at RT on an overhead tumbler.
10. Centrifuge at 1400 g, 3 min. Discard the supernatant.
11. Transfer into 1.5 ml caps and wash
    3-times with 1.5 ml 0.2% SDS/PBS (washing time: 0, 5 and 20 min)
    3-times with 1.5 ml PBS (washing time: 0, 5 and 10 min)
    3-times with 1.5 ml water (washing time: 0-1 min).

Digestion

12. Discard supernatant. Add 500 µl of 6 M Urea/PBS, 10 mM DTT.
13. Incubate at 65°C, 15 min, allow to cool down to RT.
14. Add IAA to 40 mM (45 µl of 440 mM IAA in water).
15. Incubate 15 min at 37°C on over-head tumbler.
16. Dilute the reaction to 1.5 ml with PBS, centrifuge and discard the supernatant.
17. Add a solution of 100 µl 2 M urea/PBS, 2 µl 100 mM $CaCl_2$, 2 µg trypsin.
18. Incubate o.n. at 37°C on an overhead tumbler.
19. Transfer the supernatant into 1.5 ml caps. Wash the beads 2-times with 25 µl water and combine with the supernatant.
20. Freeze the supernatant in liquid nitrogen and dry on SpeedVac.
21. For LC-MS analysis, dissolve in 110 µl 0.1 % formic acid/water and inject 100 µl on a C18 column.

Elution of the labelled fragment

22. Wash the beads from step 19. with 1 ml 0.2 % SDS/PBS.
23. Wash 2-times with 1 ml PBS and once with 1 ml water.
24. Add 300 µl 50 % acetonitrile/water.
25. Incubate 15 min at 65°C.
26. Centrifuge at 1400 g, 3 min. Transfer the supernatant in a new cap, freeze in liquid nitrogen and dry on SpeedVac.

7.1.8.3 Enzymatic digestion

Digestion of proteins was performed either in solution, on column or on beads. Digestion on beads is described in the previous section.

Digestion in solution

Digestion in solution was mostly used during the optimisation phase.

Solutions needed: 50 mM Tris, pH 8.0
6 M urea in 50 mM Tris, pH 8.0, 10 mM DTT (prepare fresh before use)
50 mM NH_4HCO_3 in water (prepare fresh daily)

440 mM IAA in 50 mM Tris, pH 8.0 (prepare fresh before use)
1 M DTT

1. Dilute the protein to 5 µM with 6 M urea.
2. Incubate at 65°C for 15 min. Allow the reaction to cool down.
3. Add 440 mM IAA to a final concentration of 40 mM.
4. Incubate at 25°C for 15 min.
5. Add 1 M DTT to a final concentration of 30 mM.
6. Incubate at 25°C for 15 min.
7. Dilute to 0.6 M urea with 50 mM NH_4HCO_3.
8. Add trypsin in 1:50 w/w ratio over protein weight.
9. Incubate at 37°C for 4 h.

Digestion on column

The protocol of protein digestion on column was developed by Wisniewski *et al.* for proteins difficult to solubilise[129]. Here, it offers the additional advantage of getting rid of unbound compound and concentrating the protein.

Solutions needed: 1 M DTT in water
2 % SDS in 100 mM Tris, pH 8.0
Buffer UA: 8 M urea in 100 mM Tris, pH 8.5 (prepare fresh daily)
Buffer UB: 8 M urea in 100 mM Tris, pH 8.0 (prepare fresh daily)
50 mM IAA in buffer UA (prepare fresh before use)
Buffer ABC: 50 mM NH_4HCO_3 in water (prepare fresh daily)

1. To 100 µl of 1 µM protein solution, add 5.9 µl 2 % SDS and 11.7 µl 1 M DTT (final concentrations: 0.1 % SDS, 100 mM DTT).
2. Incubate at 95°C for 5 min.
3. Add 382.4 µl buffer UA to the filter unit and mix with the protein solution.
4. Centrifuge at 14 000 g, 20°C for 40 min.
5. Discard the flow-through.
6. Add 200 µl buffer UA to the unit and centrifuge at 14 000 g, 20°C, 40 min. Discard the flow-through.
7. Add 100 µl IAA solution and mix at 600 rpm in a thermo-mixer at 20°C for 1 min. Incubate without mixing for 5 min.
8. Centrifuge at 14 000 g, 20°C, 30 min.
9. Add 100 µl of buffer UB and centrifuge at 14 000 g, 20°C, 40 min. Repeat once.
10. Add 100 µl of buffer ABC and centrifuge at 14 000 g, 20°C, 40 min. Transfer to a new collection tube.
11. Add 80 µl buffer ABC with trypsin in 1:10 w/w ratio over protein weight. Mix at 600 rpm for 1 min.

12. Incubate the units at 37°C o.n. in a wet chamber.
13. Centrifuge at 14 000 g, 20°C, 40 min.
14. Add 50 µl buffer UB and centrifuge the filter units at 14 000 g, 20°C, 20 min.

7.1.8.4 Data analysis

LC-MS data were opened with Compass DataAnalysis for compound detection. The compounds were then imported in BioTools and searched against the Mascot database.

Compound detection in Compass

1. Open the analysis file in Compass DataAnalysis.
2. Go to 'Edit' - 'Chromatograms...' (F7).
3. Choose Type: 'Base Peak Chromatogram' - 'Add'.
4. The Base Peak Chromatogram (BPC) signal should be similar to the Total Ion Chromatogramm of MSMS (TIC +All MSn). If not, the ion selection for MSMS failed. Check you measuring parameters and if necessary recalibrate the system.
5. Go to 'Find' - 'Parameters'. In the 'AutoMS(n) tab' set the intensity threshold and the maximum number of compounds according to your spectrum and sample. Tick 'Fragments
qualified by' and choose amino acids. Press 'OK'.
6. Select 'Find' - 'Compounds - AutoMS(n)'.
7. Export the compounds: 'File' - 'Export' - 'Compounds...'. Choose 'Compound XML File' as file type.

BioTools – Mascot search

1. Go BioTools and open your exported compounds.
2. Select 'View' - 'Activate MS/MS view'.
3. Select 'Search' - 'Mascot MS/MS ion search'.
4. Use the URL: http://www.matrixscience.com/cgi/nph-mascot.exe?1.
5. Select 'Carbamidomethyl (C)' as 'Global Modifications' and 'Oxidation (M)' as 'Variable Modifications'.
6. Choose 'SwissProt' as database and 'Trypsin' as enzyme.
7. Use a 'Mass Tol. MS' of 400 ppm and 'MS/MS Tol' of 0.5 Da. 'Charge State': '1+. 2+ and 3+'. Press 'start'.
8. Look at the results and download the hits of interest with the button 'Get Hit(s)'.
9. Check the errors on the ions mass. If you are measuring the standard sample and you notice an high deviation, it could be necessary to perform a 'Scan calibration' (p. 137).

7.1.9 Immunofluorescence

1. Sterilise the coverslips with ethanol and lay one in each well of a 6-well culture plate.
2. Dilute the poly-L-lysine solution (0.1 % w/v in H_2O, Sigma) to 0.1 mg/ml and sterilise by filtration.
3. Dispense on the coverslips and incubate for 30 min at 37°C.
4. Wash 2-times with water and dry on air.
5. Seed 150 000 H460 cells pro well in DMEM with 10 % FCS.
6. Incubate 24 h at 37°C, 5 % CO_2.
7. Remove the medium and add new medium with either
   15 µM SecinH3 in 0.4 % DMSO/DMEM
   0.4 % DMSO/DMEM
   50 nM demecolcine in DMEM
   100 nM taxol in 0.4 % DMSO/DMEM
   or DMEM
8. Incubate o.n. at 37°C, 5 % CO_2.
9. Remove medium and add new medium as in step
10. Take the coverslip, dip it in PBS and transfer it in cold (-20°C methanol).
11. Incubate at -20°C for at least 20 min.
12. Take the coverslip, dip it in PBS an transfer on parafilm. Cover directly with PBS.
13. Wash 3-times with PBS.
14. Dilute the FITC coupled DM1A antibody 1:100 in 5 % BSA/TBS. Add to the coverslips and incubate for 1 h at RT in the dark.
15. Wash 3-times with PBS and once with 0.5X PBS.
16. Dip in water and mount with VECTASHIELD mounting medium with DAPI (Vectorlabs).

7.2 Materials

7.2.1 Special chemicals and materials for synthesis

4-Bromomethyl benzoic acid	Alfa Aesar
4-(1-Azi-2,2,2-trifluoroethyl)benzoic acid (TPD-COOH)	Bachem (Q-1540)
D-Desthiobiothin	IBA (2-1000-002)
HPLC column: MultoHigh-Bio 300-5 C4, 125 × 4 mm (Column-Nr. 2512-07)	CS-Chromatographie Service (Art.-Nr. 584 1177)
HPLC system	Agilent, 1100 serie
PS3 peptide synthesiser	Protein technologies, Inc.

7.2.2 Nucleic acids

7.2.2.1 Primers

Primers were bought from Metabion deprotected and desalted. Following primers were used:

ARNO-E156K ap	TCTGGGCCTTTCCGGGTAGGCGAAAGC
ARNO-E156K sp	GCTTTCGCCTACCCGGAAAGGCCCAGA
ARNOpbr ap	GAGATGCTGGCAGCGTGAAAGAAGCGGATTTCA
ARNOpbr sp	TGAAATCCGCTTCTTTCACGCTGCCAGCATCTC
BamHI/EspI-EGFR sp	CCGTCTCGGATCCCGGAAGCGCACGCTGCG
C-FLuc-XbaI	ATGTCTAGATTACTTGCCGCCCTTCTTGGCCTTAATGAG
Cyh2-KpnI	TAGGTACCGGGCTGCTCCTGCTTCTTCTTGACTGA
EcoRI-ARNO-S7	AGGAATTCTGATATGGCCAATGAGGGCAGTAAGACCTTGCAAC
GST sp	TGTTCGAAGATCGTTTATGTCATAAAACATATTTAAAT
GST-BamHI ap	GGTTCCGCGTGGATCCCCGGG
GST-BamHI sp	GGTTCCGCGTGGATCCCCGGG
GST-EGFR ap	TGCGAATTCGAGACGTCATGCTCCAATAAATTCACTGCTTTGTGG
GST-EGFR sp	CCGTCTCGGATCCCGAAGGCGCCACATCGTTCG
KpnI-C-FLuc	ATGGTACCATGAGCGGCTACGTTAACAACCCCGA
KpnI-Grb2	CGGTACCATGGAAGCCATCGCCAAATATGACTTC
KpnI-Grb2	CGGTACCATGGAAGCCATCGCCAAATATGACTTC
KpnI-N-FLuc	ATGGTACCGAAGATGCCAAAAACATTAAGAAGGGCCCA
mmCyh3-XhoI	ATCCTCGAGCTATTTCTTATTGGCAATCCTCCTT
mmCyh3pbr ap	ACATGTTGGCCACGTGAAAAAGGAGGATTGCCAATAAGAAATAG
N-FLuc-XbaI	ATGTCTAGATTAGCCGTCCTTGTCGATGAGAGCGTT
PstI-GST sp	CACTGCAGCATGTCCCCTATACTAGGTTATTGGAAAATTAAGG
Shc-XhoI	CTCTCGAGGGATTGGAGGGCATCTTCTGGAA

7.2.2.2 Plasmids

The plasmid cloned during this thesis are listed in Table 7.3 (p. 151). The parent plasmids, primers, vectors and enzymes used are given.

7.2.3 Buffers for SDS-PAGE

4 X Separating gel buffer	Tris pH 8.8	1.5 M
	SDS	0.4 % w/v
4 X Stacking gel buffer	Tris pH 6.8	500 mM
	SDS	0.4 % w/v

5 X Running buffer	Tris pH 8.2	125 mM
	glycine	960 mM
	SDS	0.5 % w/v
6 X Loading buffer	Tris pH 6.8	49.5 mM
	DTT	600 mM
	SDS	15 % w/v
	glycerin	30 %
	bromophenolblue	

7.2.4 Buffers and materials for Western blotting

Cathode buffer	Tris pH 9.4	25 mM
	glycine	40 mM
Anode I buffer	Tris pH 10.4	300 mM
Anode II buffer	Tris pH 10.4	25 mM
10 X TBS buffer	Tris pH 7.6	200 mM
	NaCl	1.36 M
TBS-T buffer	Tris pH 7.6	20 mM
	NaCl	136 mM
	Tween-20	0.1 % v/v
Odyssey scanner		Licor
VersaDoc 5000 CCD camera		BioRad

7.2.5 Buffers and materials for BRET

DMEM	PAA
DPBS	PAA
Metafectene	Biontex
Plate reader	Mithras LB 940 (Berthold technologies)

7.2.6 Buffers and materials for SPR

NTA running buffer	HEPES, pH 7.4	10 mM
	NaCl	300 mM
	EDTA	50 µM
	Tween-20	0.005 %
NTA dispenser buffer	HEPES, pH 7.4	10 mM
	NaCl	300 mM
	EDTA	3 mM
	Tween-20	0.005 %
NTA regeneration buffer	HEPES, pH 7.4	10 mM

	NaCl	500 mM
	EDTA	0.5 M
	Tween-20	0.005 %
$NiCl_2$ solution	HEPES, pH 7.0	10 mM
	$NiCl_2$	20 mM
	NaCl	150 mM
	Tween-20	0.005 %
A-buffer	Tris, pH 7.8	50 mM
	NaCl	300 mM
	EDTA	50 µM
	Tween-20	0.005 %
ARNO buffer	Tris pH 7.4	50 mM
	NaCl	300 mM
	$MgCl_2$	3 mM
	Tween-20	0.005 %
	DMSO	0.5 %
Small molecule coupling buffer	Na_2CO_3, pH 10.5	10 mM
	DMSO	5 %
HBS-P (Biacore)	HEPES pH 7.4	10 mM
	NaCl	150 mM
	Surfactant P20	0.005 %

NB: Surfactant P20 is commonly known as Tween 20.

Both chips from Biacore (GE Healtchare Life Sciences) and Xantec (XanTec bioanalytics GmbH, Düsseldorf) were used. In the following table the notation is compared:

Coating of the SPR sensor chip	Notation	
	Biacore	Xantec
Carboxymethyldextran	CM5	CMDP
NTA derivatised carboxymethyldextran	NTA	NiP

All measurements were performed on a Biacore 3000.

7.2.7 Buffers and materials for CCMS

10 X Labelling buffer	Tris pH 7.65	500 mM
	NaCl	3 M
	$MgCl_2$	30 mM

1 X Labelling buffer	Tris pH 7.65	50 mM
	NaCl	300 mM
	$MgCl_2$	3 mM
	glycerol	10 %

Amicon Ultra-0.5 mL Centrifugal Filters	Millipore
BSA	Calbiochem
ESI ion trap	Bruker Daltonics Esquire HCT
EZ-link iodoacetyl-PEG_2-biotin	Piercenet (#21334)
HPLC system	Agilent, 1100 serie
Iodoacetamide	Sigma
Trypsin Gold, Mass Spectrometry Grade	Promega (#TB309)
Trypsin-digested BSA (CAM-modified)	New England BioLabs (#P8108S)
Tubulin (porcine)	Cytoskeleton (#T240)
ES tuning mix	Agilent (# G2421A)
3UV ultraviolet lamps	Thermo Scientific Pierce (producer: UVP)

7.2.8 Materials for immunofluorescence

Demecolcine (10 µg/ml in HBSS)	Sigma (#D1925)
Digitonin	Sigma (#D141)
poly-L-lysine solution (0.1 % w/v in H_2O)	Sigma (#P8920)
Taxol (paclitaxel)	Sigma (#T7191)
Monoclonal anti-α-tubulin, DM1A FITC conjugate	Sigma (#F2168)
Pierce Streptavidin UltraLink Resin	Thermo Scientific (#53114)
VECTASHIELD Mounting Medium with DAPI	Vectorlabs (#H-1200)

Cloned	Plasmid Template	Vector (Resistance)	Primer Sense	Primer Antisense	Enzymes
pARNOdCC-GFP2	pARNO-RLuc	pGFP2-N2 (ZEO)	EcoRI-ARNOS7	cyh2-KpnI	EcoRI, KpnI
pARNO(E156K)-GFP2		pARNO-GFP2 (ZEO)	ARNO-E156K sp	ARNO-E156K ap	(QuickChange)
pCMVTag2-ARNO(E156K)		pCMVTag2-ARNO(KAN)	ARNO-E156K sp	ARNO-E156K ap	(QuickChange)
pCMVTag2-ARNO-Sec7(E156K)		pCMVTag2-ARNO-Sec7	ARNO-E156K sp	ARNO-E156K ap	(QuickChange)
pASK-IBA-Grb2	cDNA	pASK-IBA 43 plus (Amp)	KpnI-Grb2	Grb2-XhoI	KpnI, XhoI
pASK45-Cyt2-ΔCC	pCMVTag2-ARNOdCC	pASK-IBA 45 plus (Amp)	none[a]		EcoRI, HindIII
pASK45-Cyt2-ΔPH	pCMVTag2-ARNO-CC-Sec7	pASK-IBA 45 plus (Amp)	none[a]		EcoRI, HindIII
pASK45-Cyt2-Full	pCMVTag2-ARNO	pASK-IBA 45 plus (Amp)	none[a]		EcoRI, HindIII
pASK45-Cyt2-Full(E156K)	pCMVTag2-ARNO(E156K)	pASK-IBA 45 plus (Amp)	none[a]		EcoRI, HindIII
pASK45-Cyt2-Sec7	pCMVTag2-ARNO-Sec7	pASK-IBA 45 plus (Amp)	none[a]		EcoRI, HindIII
pASK45-Cyt2-Sec7(E156K)	pCMVTag2-ARNO-Sec7(E156K)	pASK-IBA 45 plus (Amp)	none[a]		EcoRI, HindIII
pCMVneo-EGFR-NLuc	pGL4.12[luc2CP]	pEGFR-RLuc (KAN)[b]	KpnI-N-FLuc	N-FLuc-XbaI	KpnI, XbaI
pCMVzeo-EGFR-CLuc	pGL4.12[luc2CP]	pEGFR-GFP2 (ZEO)[b]	KpnI-C-FLuc	C-FLuc-XbaI	KpnI, XbaI
pCMVzeo-EGFR-FLuc	pGL4.12[luc2CP]	pEGFR-GFP2 (ZEO)[b]	KpnI-N-FLuc	C-FLuc-XbaI	KpnI, XbaI
pET11-ARNOpbr		pET11-ARNO (Amp)	ARNOpbr sp	ARNOpbr ap	(QuickChange)
pET15b-mmCyh3pbr		pET15b-mmCyh3 (Amp)	mmCyh3pbr sp	mmCyh3pbr ap	(QuickChange)
pGEX-GST-EGFR	pEGFR-RLuc	pGEX-2T (Amp)	BamHI/EspI-EGFR	GST-EGFR ap	Esp3I (insert) BamHI, EcoRI (vector)
pGST-GFP2	pGEX-2T	pGFP2-N2 (ZEO)	PstI-GST sp	GST-BamHI ap	PstI, BamHI
pGST-RLuc	pGEX-2T	pRLuc-N2 (KAN)	PstI-GST sp	GST-BamHI ap	PstI, BamHI

Table 7.3: Plasmid constructs cloned - ZEO: Zeocin; KAN: kanamycin; Amp: Ampicillin.

[a]The inserts were obtained by digestion of the template plasmid.

[b]XbaI is methylation sensitive. The vectors were purified from Dam$^-$ E.Coli.

Acknowledgements

Keep a cell alone, nothing will happen. Give it signals, it will grow. Allow it to communicate with various different cells, you will get a fully functional organism.

The same is true for a PhD student. Therefore I would like to gratefully thank not only my colleagues who helped me to scientifically grow, but also my family, friends and all those persons who in the most disparate ways created the right environment for my development. My special thanks go to (in alphabetical order):

- Alexander, Adrian, Monika and Ulf and the Lebenshilfe Rhein-Sieg e.V., for offering me a perfect recreation opportunity. A&A: your sincere smiling when seeing me, always gave me energy for a week.

- Anke, for our friendship, for explaining me the difference between a sheet full of spots and a Western blot and for our very diversified experiences in and outside the lab.

- Christina, who always realised when I needed some encouragement and for her gift of surprising me.

- Christine, for the paper-coffees and the coffees without papers and for offering regular highlights to my writing phase.

- Damian, for his invaluable help in synthesis design, HPLC, MS and especially for his support in the difficult times.

- my boyfriend Daniel, for the wholehearted support through the ups and downs and for always trying to show me the positive face of the situation.

- the members of my distributed proofreader team, who accepted to read and comment some short or long parts of this work: Toni, Christine, Daniel, Damian, Martin, Gen and Jeff.

- Esteban, for his active participation to the SPR adventure and all the side effects.

- my supervisor, Prof. Dr. Famulok, for showing me the light at the end of the tunnel.

- the whole Famulok and Mayer labs for their help, friendliness and the availability for a chat when necessary.

- PD Dr. Franken, for donating some hours of his time, which saved weeks of mine.

- all my friends, who never abandoned me, even if I didn't contact them for months.

- Gen and Christian, for the good neighbourhood and the motivating time spent together.

- Giosiana, for always finding some time for me during my flying visits in Ticino.

- my grandparents, uncles, aunts, cousins, ... for their support and friendship.

- the GRK804 members, for the good time spent together and the interesting scientific (and less scientific) activities.

- Heike, who volunteered as supporter for immunofluorescence and always had some good words and compound solutions for me.

- Jan, for taming with me the LC-MS instrument.

- Jeff, for its help in chemical questions and for the long evening chats.

- Justina for preventing our sinking in chaos and for the appreciated Korean catering.

- Martin, for sharing with me his microscopy knowledge and for its positiveness.

- Prof. Dr. Mayer, for his realistic and pragmatic vision of life and for accepting the role of 2nd referee.

- Prof. Dr. Piel for accepting within 2 minutes and 32 seconds to be my 3rd referee.

- Dr. Rich and Dr. Myszka, for their unconventional reviews and their commitment in establishing quality standards for SPR.

- the Roche Research Foundation and the GRK804, for financial support.

- my parents, Paola and Silvano for their constant encouragement, for adapting their plans to my holidays and for doing their best to understand what I'm doing in the lab.

- Sven, for his straightforward help in all administrative issues and for his engagement in the reanimation of the MS-computer.

- my brother Tazio, for his helpfulness at every time and for keeping me (and my cellar) up-to-date on the (agricultural) technical advances.

- Toni, for the patience in answering my questions, the unlimited tipps&tricks and his perseverance in looking for new solutions.

- Volkmar, for the help in molecular cloning and protein purification.

- the Zheng Lab, for turning my thoughts into music.

Bibliography

[1] Bos, J. L., Rehmann, H. & Wittinghofer, A. GEFs and GAPs: critical elements in the control of small G proteins. *Cell* **129**, 865–877 (2007).

[2] Vigil, D., Cherfils, J., Rossman, K. L. & Der, C. J. Ras superfamily GEFs and GAPs: validated and tractable targets for cancer therapy? *Nat. Rev. Cancer.* **10**, 842–857 (2010).

[3] Kahn, R. A. & Gilman, A. G. Purification of a protein cofactor required for ADP-ribosylation of the stimulatory regulatory component of adenylate cyclase by cholera toxin. *J. Biol. Chem.* **259**, 6228–6234 (1984).

[4] Kahn, R. A. & Gilman, A. G. The protein cofactor necessary for ADP-ribosylation of Gs by cholera toxin is itself a GTP binding protein. *J. Biol. Chem.* **261**, 7906–7911 (1986).

[5] Casanova, J. E. Regulation of Arf activation: the Sec7 family of guanine nucleotide exchange factors. *Traffic* **8**, 1476–1485 (2007).

[6] Jackson, C. L. & Casanova, J. E. Turning on ARF: the Sec7 family of guanine-nucleotide-exchange factors. *Trends Cell Biol.* **10**, 60 – 67 (2000).

[7] Gillingham, A. K. & Munro, S. The small G proteins of the Arf family and their regulators. *Annu. Rev. Cell Dev. Biol.* **23**, 579–611 (2007).

[8] Pasqualato, S., Renault, L. & Cherfils, J. Arf, Arl, Arp and Sar proteins: a family of GTP-binding proteins with a structural device for 'front–back' communication. *EMBO reports* **3**, 1035–1041 (2002).

[9] Renault, L., Guibert, B. & Cherfils, J. Structural snapshots of the mechanism and inhibition of a guanine nucleotide exchange factor. *Nature* **426**, 525–530 (2003).

[10] Goldberg, J. Structural basis for activation of Arf GTPase: mechanisms of guanine nucleotide exchange and GTP-myristoyl switching. *Cell* **95**, 237 – 248 (1998).

[11] Cherfils, J. *et al.* Structure of the Sec7 domain of the Arf exchange factor ARNO. *Nature* **392**, 101–105 (1998).

[12] Beraud-Dufour, S. *et al.* A glutamic finger in the guanine nucleotide exchange factor ARNO displaces Mg^{2+} and the β-phosphate to destabilize GDP on ARF1. *EMBO J.* **17**, 3651–3659 (1998).

[13] Helms, J. B. & Rothman, J. E. Inhibition by brefeldin A of a Golgi membrane enzyme that catalyses exchange of guanine nucleotide bound to ARF. *Nature* **360**, 352–354 (1992).

[14] Donaldson, J. G., Finazzi, D. & Klausner, R. D. Brefeldin A inhibits Golgi membrane-catalysed exchange of guanine nucleotide onto ARF protein. *Nature* **360**, 350–352 (1992).

[15] Jackson, C. L. Brefeldin A revealing the fundamental principles governing membrane dynamics and protein transport. *Subcell. Biochem* **34**, 233–272 (2000).

[16] Viaud, J. *et al.* Structure-based discovery of an inhibitor of Arf activation by Sec7 domains through targeting of protein–protein complexes. *PNAS* **104**, 10370–10375 (2007).

[17] Hafner, M. *et al.* Inhibition of cytohesins by SecinH3 leads to hepatic insulin resistance. *Nature* **444**, 941–944 (2006).

[18] Kolanus, W. Guanine nucleotide exchange factors of the cytohesin family and their roles in signal transduction. *Immunol. Rev.* **218**, 102–113 (2007).

[19] DiNitto, J. P. *et al.* Structural basis and mechanism of autoregulation in 3-phosphoinositide-dependent Grp1 family Arf GTPase exchange factors. *Mol. Cell* **28**, 569–583 (2007).

[20] Santy, L. C., Frank, S. R., Hatfield, J. C. & Casanova, J. E. Regulation of ARNO nucleotide exchange by a PH domain electrostatic switch. *Curr. Biol.* **9**, 1173–1176 (1999).

[21] Dierks, H., Kolanus, J. & Kolanus, W. Actin cytoskeletal association of cytohesin-1 is regulated by specific phosphorylation of its carboxyl-terminal polybasic domain. *J. Biol. Chem.* **276**, 37472–37481 (2001).

[22] Kolanus, W. *et al.* αLβ2 integrin/LFA-1 binding to ICAM-1 induced by cytohesin-1, a cytoplasmic regulatory molecule. *Cell* **86**, 233–242 (1996).

[23] Kliche, S. *et al.* Signaling by human herpesvirus 8 kaposin A through direct membrane recruitment of cytohesin-1. *Mol. Cell* **7**, 833–843 (2001).

[24] Hurtado-Lorenzo, A. *et al.* V-ATPase interacts with ARNO and Arf6 in early endosomes and regulates the protein degradative pathway. *Nat. Cell Biol.* **8**, 124–136 (2006).

[25] Gsandtner, I. *et al.* Heterotrimeric G protein-independent signaling of a G protein-coupled receptor. Direct binding of ARNO/cytohesin-2 to the carboxyl terminus of the A_{2A} adenosine receptor is necessary for sustained activation of the ERK/MAP kinase pathway. *J. Biol. Chem.* **280**, 31898–31905 (2005).

[26] Stalder, D. *et al.* Kinetic studies of the Arf activator Arno on model membranes in the presence of Arf effectors suggest control by a positive feedback loop. *J. Biol. Chem.* **286**, 3873–3883 (2011).

[27] Jackson, C. Diabetes: Kicking off the insulin cascade. *Nature* **444**, 833–834 (2006).

[28] Li, C.-C. *et al.* ARL4D recruits cytohesin-2/ARNO to modulate actin remodeling. *Mol. Biol. Cell* **18**, 4420–4437 (2007).

[29] Várnai, P. *et al.* Selective cellular effects of overexpressed pleckstrin-homology domains that recognize PtdIns(3,4,5)P3 suggest their interaction with protein binding partners. *Journal of Cell Science* **118**, 4879–4888 (2005).

[30] Hofmann, I., Thompson, A., Sanderson, C. & Munro, S. The Arl4 family of small G proteins can recruit the cytohesin Arf6 exchange factors to the plasma membrane. *Curr. Biol.* **17**, 711 – 716 (2007).

[31] Cohen, L. A. *et al.* Active Arf6 recruits ARNO/cytohesin GEFs to the PM by binding their PH domains. *Mol. Biol. Cell* **18**, 2244 (2007).

[32] Mayer, G. *et al.* Controlling small guanine-nucleotide-exchange factor function through cytoplasmic RNA intramers. *PNAS* **98**, 4961–4965 (2001).

[33] Hafner, M. *et al.* Displacement of protein-bound aptamers with small molecules screened by fluorescence polarization. *Nat. Protocols* **3**, 579–587 (2008).

[34] Stumpfe, D. *et al.* Targeting multifunctional proteins by virtual screening: structurally diverse cytohesin inhibitors with differentiated biological functions. *ACS Chem. Biol.* **5**, 839–849 (2010).

[35] Fuss, B., Becker, T., Zinke, I. & Hoch, M. The cytohesin Steppke is essential for insulin signalling in Drosophila. *Nature* **444**, 945–948 (2006).

[36] Bill, A. *Cytohesine beeinflussen die wachstumsfaktorabhängige Proliferation von Krebszellen.* Master's thesis, Rheinischen Friedrich-Wilhelms-Universität Bonn (2007).

[37] Tan, P. K. *et al.* Monitoring interactions between receptor tyrosine kinases and their downstream effector proteins in living cells using bioluminescence resonance energy transfer. *Mol. Pharmacol.* **72**, 1440–1446 (2007).

[38] Schlessinger, J. Cell signaling by receptor tyrosine kinases. *Cell* **103**, 211–225 (2000).

[39] Blume-Jensen, P. & Hunter, T. Oncogenic kinase signalling. *Nature* **411**, 355–365 (2001).

[40] Yarden, Y. & Sliwkowski, M. X. Untangling the ErbB signalling network. *Nat. Rev. Mol. Cell Biol.* **2**, 127–137 (2001).

[41] Zhang, X., Gureasko, J., Shen, K., Cole, P. A. & Kuriyan, J. An Allosteric Mechanism for Activation of the Kinase Domain of Epidermal Growth Factor Receptor. *Cell* **125**, 1137–1149 (2006).

[42] Chung, I. *et al.* Spatial control of EGF receptor activation by reversible dimerization on living cells. *Nature* **464**, 783–787 (2010).

[43] Bose, R. & Zhang, X. The ErbB kinase domain: Structural perspectives into kinase activation and inhibition. *Exp. Cell Res.* **315**, 649 – 658 (2009).

[44] Zhang, X. *et al.* Inhibition of the EGF receptor by binding of MIG6 to an activating kinase domain interface. *Nature* **450**, 741–744 (2007).

[45] Xu, Y., Piston, D. W. & Johnson, C. H. A bioluminescence resonance energy transfer (BRET) system: Application to interacting circadian clock proteins. *PNAS* **96**, 151–156 (1999).

[46] Pfleger, K. D. G. & Eidne, K. A. Illuminating insights into protein-protein interactions using bioluminescence resonance energy transfer (BRET). *Nat. Methods* **3**, 165–174 (2006).

[47] Wu, P. G. & Brand, L. Resonance energy transfer: methods and applications. *Anal. Biochem.* **218**, 1 – 13 (1994).

[48] Ramsay, D., Kellett, E., McVey, M., Rees, S. & Milligan, G. Homo-and hetero-oligomeric interactions between G-protein-coupled receptors in living cells monitored by two variants of bioluminescence resonance energy transfer (BRET): hetero-oligomers between receptor subtypes form more efficiently than between less closely related sequences. *Biochem. J.* **365**, 429 (2002).

[49] Pfleger, K. D. G., Seeber, R. M. & Eidne, K. A. Bioluminescence resonance energy transfer (BRET) for the real-time detection of protein-protein interactions. *Nat. Protoc.* **1**, 337–345 (2006).

[50] Bill, A. *et al.* Cytohesins are cytoplasmic ErbB receptor activators. *Cell* **143**, 201 – 211 (2010).

[51] Ogiso, H. *et al.* Crystal structure of the complex of human epidermal growth factor and receptor extracellular domains. *Cell* **110**, 775–787 (2002).

[52] Wang, H. *et al.* Identification of an exon 4-deletion variant of epidermal growth factor receptor with increased metastasis-promoting capacity. *Neoplasia* **13**, 461–71 (2011).

[53] Venkateswarlu, K., Oatey, P. B., Tavar, J. M. & Cullen, P. J. Insulin-dependent translocation of ARNO to the plasma membrane of adipocytes requires phosphatidylinositol 3-kinase. *Curr. Biol.* **8**, 463 – 466 (1998).

[54] Venkateswarlu, K. Interaction protein for cytohesin exchange factors 1 (IPCEF1) binds cytohesin 2 and modifies its activity. *J. Biol. Chem.* **278**, 43460 (2003).

[55] Venkateswarlu, K. Analysis of the interaction between cytohesin 2 and IPCEF1. *Meth. Enzymol.* **404**, 252 – 266 (2005).

[56] Torii, T. *et al.* Cytohesin-2/ARNO, through its interaction with focal adhesion adaptor protein paxillin, regulates preadipocyte migration via the downstream activation of Arf6. *J. Biol. Chem.* **285**, 24270–24281 (2010).

[57] Oatey, P. B. *et al.* Confocal imaging of the subcellular distribution of phosphatidylinositol 3,4,5-trisphosphate in insulin-and PDGF-stimulated 3T3-L1 adipocytes. *Biochem. J.* **344**, 511 (1999).

[58] Lauritzen, H. P. M. M., Ploug, T., Prats, C., Tavaré, J. & Galbo, H. Imaging of insulin signaling in skeletal muscle of living mice shows major role of T-tubules. *Diabetes* **55**, 1300 (2006).

[59] DiNitto, J. P., Lee, M.-T., Malaby, A. W. & Lambright, D. G. Specificity and membrane partitioning of Grsp1 signaling complexes with Grp1 family Arf exchange factors. *Biochemistry* **49**, 6083–6092 (2010).

[60] Yang, K. S., Ilagan, M., Xenia, G., Piwnica-Worms, D. & Pike, L. J. Luciferase fragment complementation imaging of conformational changes in the epidermal growth factor receptor. *J. Biol. Chem.* **284**, 7474 (2009).

[61] Homola, J. Electromagnetic theory of surface plasmons. In Wolfbeis, O. S. & Homola, J. (eds.) *Surface plasmon resonance based sensors*, vol. 4 of *Springer series on chemical sensors and biosensors*, 3–44 (Springer Berlin Heidelberg, 2006).

[62] GE Healthcare Bio-Sciences AB. *Biacore - Sensor surface handbook* (2005).

[63] Homola, J. & Piliarik, M. Surface plasmon resonance (SPR) sensors. In Wolfbeis, O. S. & Homola, J. (eds.) *Surface plasmon resonance based sensors*, vol. 4 of *Springer series on chemical sensors and biosensors*, 45–67 (Springer Berlin Heidelberg, 2006).

[64] Biacore AB. *BIACORE 3000 - Instrument handbook* (1999).

[65] Rich, R. L. & Myszka, D. G. Survey of the year 2007 commercial optical biosensor literature. *J. Mol. Recognit.* **21**, 355–400 (2008).

[66] Myszka, D. G. Improving biosensor analysis. *J. Mol. Recognit.* **12**, 279–284 (1999).

[67] Löfås, S. & Mcwhirter, A. The art of immobilization for SPR sensors. In Wolfbeis, O. S. & Homola, J. (eds.) *Surface plasmon resonance based sensors*, vol. 4 of *Springer series on chemical sensors and biosensors*, 117–151 (Springer Berlin Heidelberg, 2006).

[68] Hochuli, E., Bannwarth, W., Dobeli, H., Gentz, R. & Stuber, D. Genetic approach to facilitate purification of recombinant proteins with a novel metal chelate adsorbent. *Nat Biotech* **6**, 1321–1325 (1988).

[69] Štěpánek, J., Vaisocherová, H. & Piliarik, M. Molecular interactions in SPR sensors. In Wolfbeis, O. S. & Homola, J. (eds.) *Surface plasmon resonance based sensors*, vol. 4 of *Springer series on chemical sensors and biosensors*, 69–91 (Springer Berlin Heidelberg, 2006).

[70] Navratilova, I. & Myszka, D. G. Investigating biomolecular interactions and binding properties using SPR biosensors. In Wolfbeis, O. S. & Homola, J. (eds.) *Surface plasmon resonance based sensors*, vol. 4 of *Springer series on chemical sensors and biosensors*, 155–176 (Springer Berlin Heidelberg, 2006).

[71] Vorup-Jensen, T. Coping with complexity (in macromolecular interactions)-a comment on Rebecca L. Rich's and David G. Myszka's Grading the commercial optical biosensor literature-Class of 2008: The Mighty Binders. *J. Mol. Recognit.* **23**, 389391 (2010).

[72] Schmidt, T. G. & Skerra, A. The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins. *Nat. Protocols* **2**, 1528–1535 (2007).

[73] Marchetti, L. *High throughput screening for the identification of small molecules that interfere with RNAi mechanism.* Master's thesis, Università degli Studi di Padova (2004).

[74] Albertoni, B. *Synthesis and analysis of inhibitors of the binding of siRNA to the PAZ domain.* Master's thesis, ETH Zürich (2007).

[75] Willard, F. S. & Siderovski, D. P. Covalent immobilization of histidine-tagged proteins for surface plasmon resonance. *Anal. Biochem.* **353**, 147 – 149 (2006).

[76] Gutierrez, E. *Binding study of a novel cytohesin-2 small molecule inhibitor: establishment of small molecule-protein interaction analysis.* Master's thesis, Rheinischen Friedrich-Wilhelms-Universität Bonn (2010).

[77] Bi, X., Schmitz, A., Hayallah, A., Song, J. & Famulok, M. Affinity-based labeling of cytohesins with a bifunctional SecinH3 photoaffinity probe. *Angew. Chem. Int. Ed. Engl.* **47**, 9565–9568 (2008).

[78] Grüne, I. *Identifizierung und in vitro Charakterisierung eines neuen Inhibitors der Sec7-Domäne der Cytohesin-Familie.* Ph.D. thesis, Rheinischen Friedrich-Wilhelms-Universität Bonn (2006).

[79] Ketley, J. N., Habig, W. & Jakoby, W. Binding of nonsubstrate ligands to the glutathione S-transferases. *J. Biol. Chem.* **250**, 8670 (1975).

[80] Kaplowitz, N. Physiological significance of glutathione S-transferases. *Am. J. Physiol. Gastrointest. Liver Physiol.* **239**, 439 (1980).

[81] Litwack, G., Ketterer, B. & Arias, I. Ligandin: a hepatic protein which binds steroids, bilirubin, carcinogens and a number of exogenous organic anions. *Nature* **234**, 466–467 (1971).

[82] Sleno, L. & Emili, A. Proteomic methods for drug target discovery. *Curr. Opin. Chem. Biol.* **12**, 46 – 54 (2008).

[83] Liu, Y., Patricelli, M. & Cravatt, B. Activity-based protein profiling: the serine hydrolases. *PNAS* **96**, 14694 (1999).

[84] Cravatt, B. F. & Sorensen, E. J. Chemical strategies for the global analysis of protein function. *Curr. Opin. Chem. Biol.* **4**, 663–668 (2000).

[85] Jeffery, D. A. & Bogyo, M. Chemical proteomics and its application to drug discovery. *Curr. Opin. Biotechnol.* **14**, 87–95 (2003).

[86] Hagenstein, M. C. *et al.* Affinity-based tagging of protein families with reversible inhibitors: a concept for functional proteomics. *Angew. Chem. Int. Ed. Engl.* **42**, 5635–5638 (2003).

[87] Saghatelian, A., Jessani, N., Joseph, A., Humphrey, M. & Cravatt, B. F. Activity-based probes for the proteomic profiling of metalloproteases. *PNAS* **101**, 10000–10005 (2004).

[88] Chan, E. W. S., Chattopadhaya, S., Panicker, R. C., Huang, X. & Yao, S. Q. Developing photoactive affinity probes for proteomic profiling: hydroxamate-based probes for metalloproteases. *JACS* **126**, 14435–14446 (2004).

[89] Adam, G. C., Burbaum, J., Kozarich, J. W., Patricelli, M. P. & Cravatt, B. F. Mapping enzyme active sites in complex proteomes. *JACS* **126**, 1363–1368 (2004).

[90] Speers, A. E. & Cravatt, B. F. A tandem orthogonal proteolysis strategy for high-content chemical proteomics. *JACS* **127**, 10018–10019 (2005).

[91] Okerberg, E. S. *et al.* High-resolution functional proteomics by active-site peptide profiling. *PNAS* **102**, 4996 (2005).

[92] Weerapana, E., Speers, A. E. & Cravatt, B. F. Tandem orthogonal proteolysis-activity-based protein profiling (TOP-ABPP) – a general method for mapping sites of probe modification in proteomes. *Nat. Protoc.* **2**, 1414 (2007).

[93] Köster, H. *et al.* Capture compound mass spectrometry: a technology for the investigation of small molecule protein interactions. *Assay Drug Dev Technol* **5**, 381–390 (2007).

[94] McLafferty, F. W. A century of progress in molecular mass spectrometry. *Annu Rev Anal Chem* **4**, null (2011).

[95] Fenn, J., Mann, M., Meng, C., Wong, S. & Whitehouse, C. Electrospray ionization for mass spectrometry of large biomolecules. *Science* **246**, 64–71 (1989).

[96] Karas, M. & Hillenkamp, F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal. Chem.* **60**, 2299–2301 (1988).

[97] Tanaka, K. *et al.* Protein and polymer analyses up to m/z 100,000 by laser ionization time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* **2**, 151–153 (1988).

[98] Hernandez, P., Müller, M. & Appel, R. D. Automated protein identification by tandem mass spectrometry: Issues and strategies. *Mass Spectrom Rev* **25**, 235–254 (2006).

[99] Markides, K. & Gräslund, A. Mass spectrometry (MS) and nuclear magnetic resonance (NMR) applied to biological macromolecules. *Advanced information on the Nobel Prize in Chemistry 2002* (2002).

[100] Steen, H. & Mann, M. The abc's (and xyz's) of peptide sequencing. *Nat. Rev. Mol. Cell Biol.* **5**, 699–711 (2004).

[101] Biemann, K. Mass spectrometry of peptides and proteins. *Annu. Rev. Biochem.* **61**, 977–1010 (1992).

[102] Roepstorff, P. & Fohlman, J. Letter to the editors. *Biol. Mass Spectrom.* **11**, 601–601 (1984).

[103] Frank, A. M., Savitski, M. M., Nielsen, M. L., Zubarev, R. A. & Pevzner, P. A. De novo peptide sequencing and identification with precision mass spectrometry. *J. Proteome Res.* **6**, 114–123 (2007).

[104] Mann, M. & Wilm, M. Error-tolerant identification of peptides in sequence databases by peptide sequence tags. *Anal. Chem.* **66**, 4390–4399 (1994).

[105] Kelleher, N. L. *et al.* Top down versus bottom up protein characterization by tandem high-resolution mass spectrometry. *JACS* **121**, 806–812 (1999).

[106] Han, X., Aslanian, A. & III, J. R. Y. Mass spectrometry for proteomics. *Curr. Opin. Chem. Biol.* **12**, 483 – 490 (2008).

[107] Armirotti, A. & Damonte, G. Achievements and perspectives of top-down proteomics. *Proteomics* **10**, 3566–3576 (2010).

[108] Olsen, J. V., Ong, S.-E. & Mann, M. Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Mol. Cell Proteomics* **3**, 608–614 (2004).

[109] Baldwin, M. A. Protein identification by mass spectrometry. *Mol. Cell Proteomics* **3**, 1–9 (2004).

[110] Henzel, W. J. *et al.* Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases. *PNAS* **90**, 5011–5015 (1993).

[111] James, P., Quadroni, M., Carafoli, E. & Gonnet, G. Protein identification by mass profile fingerprinting. *Biochem. Biophys. Res. Commun.* **195**, 58 – 64 (1993).

[112] Mann, M., Højrup, P. & Roepstorff, P. Use of mass spectrometric molecular weight information to identify proteins in sequence databases. *Biol. Mass Spectrom.* **22**, 338–345 (1993).

[113] Pappin, D. J. C., Hojrup, P. & Bleasby, A. J. Rapid identification of proteins by peptide-mass fingerprinting. *Curr. Biol.* **3**, 327 – 332 (1993).

[114] Yates, J. R., Speicher, S., Griffin, P. R. & Hunkapiller, T. Peptide mass maps: a highly informative approach to protein identification. *Anal. Biochem.* **214**, 397 – 408 (1993).

[115] Eng, J. K., McCormack, A. L. & III, J. R. Y. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom.* **5**, 976 – 989 (1994).

[116] Brunner, J. New photolabeling and crosslinking methods. *Annu. Rev. Biochem.* **62**, 483–514 (1993).

[117] Brunner, J., Senn, H. & Richards, F. 3-Trifluoromethyl-3-phenyldiazirine. A new carbene generating group for photolabeling reagents. *J. Biol. Chem.* **255**, 3313–3318 (1980).

[118] Blencowe, A. & Hayes, W. Development and application of diazirines in biological and synthetic macromolecular systems. *Soft Matter* **1**, 178–205 (2005).

[119] Hashimoto, M. & Hatanaka, Y. Recent progress in diazirine-based photoaffinity labeling. *European Journal of Organic Chemistry* **2008**, 2513–2523 (2008).

[120] Hatanaka, Y., Hashimoto, M., Kurihara, H., Nakayama, H. & Kanaoka, Y. A novel family of aromatic diazirines for photoaffinity labeling. *J. Org. Chem.* **59**, 383–387 (1994).

[121] Tanaka, Y., Bond, M. R. & Kohler, J. J. Photocrosslinkers illuminate interactions in living cells. *Mol Biosyst.* **4**, 473–480 (2008).

[122] Kotzyba-Hibert, F., Kapfer, I. & Goeldner, M. Recent trends in photoaffinity labeling. *Angew. Chem. Int. Ed. Engl.* **34**, 1296–1312 (1995).

[123] Hirsch, J. D. *et al.* Easily reversible desthiobiotin binding to streptavidin, avidin, and other biotin-binding proteins: uses for protein labeling, detection, and isolation. *Anal. Biochem.* **308**, 343–357 (2002).

[124] Rosell, C. M., Terreni, M., Fernandez-Lafuente, R. & Guisan, J. M. A criterion for the selection of monophasic solvents for enzymatic synthesis. *Enzyme Microb. Technol.* **23**, 64 – 69 (1998).

[125] Park, D.-W., Kim, H.-S., Jung, J.-K., Haam, S. & Kim, W.-S. Enzymatic synthesis of alkylglucosides by amphiphilic phase enzyme reaction. *Biotechnol. Lett.* **22**, 951–956 (2000).

[126] Kumar, S. & Hein, G. E. Structural and steric specificity of α-chloroketones as inhibitors of α-chymotrypsin. *Biochimica et Biophysica Acta (BBA) - Enzymology* **206**, 404 – 413 (1970).

[127] Shaw, W. H. R. & Bordeaux, J. J. The decomposition of urea in aqueous media. *JACS* **77**, 4729–4733 (1955).

[128] Hagel, P., Gerding, J. J. T., Fieggen, W. & Bloemendal, H. Cyanate formation in solutions of urea: I. Calculation of cyanate concentrations at different temperature and pH. *Biochimica et Biophysica Acta (BBA) - Protein Structure* **243**, 366 – 373 (1971).

[129] Wiśniewski, J. R., Zougman, A., Nagaraj, N. & Mann, M. Universal sample preparation method for proteome analysis. *Nat. Methods* **6**, 359–62 (2009).

[130] IonSource. http://www.ionsource.com/Card/carbam/carbam.htm (17 March 2011).

[131] Kim, S. C. *et al.* A clean, more efficient method for in-solution digestion of protein mixtures without detergent or urea. *J. Proteome Res.* **5**, 3446–3452 (2006).

[132] Perkins, D. N., Pappin, D. J. C., Creasy, D. M. & Cottrell, J. S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20**, 3551–3567 (1999).

[133] Plutner, H., Davidson, H. W., Saraste, J. & Balch, W. E. Morphological analysis of protein transport from the ER to Golgi membranes in digitonin-permeabilized cells: role of the P58 containing compartment. *J. Cell Biol.* **119**, 1097–1116 (1992).

[134] Salisbury, C. M. & Cravatt, B. F. Activity-based probes for proteomic profiling of histone deacetylase complexes. *PNAS* **104**, 1171–1176 (2007).

[135] Qiu, W.-W., Xu, J., Li, J.-Y., Li, J. & Nan, F.-J. Activity-based protein profiling for type I methionine aminopeptidase by using photo-affinity trimodular probes. *ChemBioChem* **8**, 1351–1358 (2007).

[136] Kunzmann, M. H., Staub, I., Böttcher, T. & Sieber, S. A. Protein reactivity of natural product-derived γ-butyrolactones. *Biochemistry* **50**, 910–916 (2011).

[137] Howard, J. & Hyman, A. A. Dynamics and mechanics of the microtubule plus end. *Nature* **422**, 753–758 (2003).

[138] Nogales, E. Structural insights into microtubule function. *Annu. Rev. Biochem.* **69**, 277–302 (2000).

[139] Thyberg, J. & Moskalewski, S. Role of microtubules in the organization of the Golgi complex. *Exp. Cell Res.* **246**, 263 – 279 (1999).

[140] Mitchison, T. & Kirschner, M. Dynamic instability of microtubule growth. *Nature* **312**, 237–242 (1984).

[141] Nogales, E. & Wang, H.-W. Structural intermediates in microtubule assembly and disassembly: how and why? *Curr. Opin. Cell Biol.* **18**, 179 – 184 (2006).

[142] Nogales, E., Wolf, S. G. & Downing, K. H. Structure of the [alpha][beta] tubulin dimer by electron crystallography. *Nature* **391**, 199–203 (1998).

[143] Akhmanova, A. & Steinmetz, M. O. Tracking the ends: a dynamic protein network controls the fate of microtubule tips. *Nat. Rev. Mol. Cell Biol.* **9**, 309–322 (2008).

[144] Howard, J. & Hyman, A. A. Microtubule polymerases and depolymerases. *Curr. Opin. Cell Biol.* **19**, 31 – 35 (2007).

[145] Westermann, S. & Weber, K. Post-translational modifications regulate microtubule function. *Nat. Rev. Mol. Cell Biol.* **4**, 938–948 (2003).

[146] Hammond, J. W., Cai, D. & Verhey, K. J. Tubulin modifications and their cellular functions. *Curr. Opin. Cell Biol.* **20**, 71 – 76 (2008).

[147] Risinger, A. L., Giles, F. J. & Mooberry, S. L. Microtubule dynamics as a target in oncology. *Cancer Treat. Rev.* **35**, 255 – 261 (2009).

[148] Rao, S. *et al.* Characterization of the taxol binding site on the microtubule. *J. Biol. Chem.* **274**, 37990–37994 (1999).

[149] Kolb, H. C., Finn, M. G. & Sharpless, K. B. Click chemistry: diverse chemical function from a few good reactions. *Angew. Chem. Int. Ed. Engl.* **40**, 2004–2021 (2001).

[150] Zhang, Y. *et al.* Carbohydrate-protein interactions by "clicked" carbohydrate self-assembled monolayers. *Anal. Chem.* **78**, 2001–2008 (2006).

[151] Wienken, C. J., Baaske, P., Rothbauer, U., Braun, D. & Duhr, S. Protein-binding assays in biological liquids using microscale thermophoresis. *Nat Commun* **1**, 100– (2010).

[152] Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248 – 254 (1976).

[153] Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685 (1970).

[154] Kyhse-Andersen, J. Electroblotting of multiple gels: a simple apparatus without buffer tank for rapid transfer of proteins from polycrylamide to nitrocellulose. *J. Biochem. Biophys. Methods* **10**, 203 – 209 (1984).

[155] Galvani, M., Hamdan, M., Herbert, B. & Righetti, P. G. Alkylation kinetics of proteins in preparation for two-dimensional maps: A matrix assisted laser desorption/ionization-mass spectrometry investigation. *Electrophoresis* **22**, 2058–2065 (2001).