Protein-protein interaction fine-tuning by estrogen receptor post-translational modification
Tharun, I.M.

DOI:
10.6100/IR782931

Published: 06/11/2014

Document Version
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:
• A submitted manuscript is the author’s version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher’s website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Protein-protein interaction fine-tuning by estrogen receptor post-translational modification

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Technische Universiteit Eindhoven, op gezag van de rector magnificus prof.dr.ir. C.J. van Duijn, voor een commissie aangewezen door het College voor Promoties, in het openbaar te verdedigen op donderdag 06 november 2014 om 16:00 uur

door

Inga Maria Tharun

geboren te Hamburg, Duitsland
Dit proefschrift is goedgekeurd door de promotoren en de samenstelling van de promotiecommissie is als volgt:

voorzitter: prof.dr. P.A.J. Hilbers
1e promotor: prof.dr.ir. L. Brunsveld
copromotor: dr. C. Ottmann
leden: prof.dr. M. Kaiser (Universität Duisburg-Essen)
       PD dr. S.A. Dames (TU München)
       prof.dr. E.W. Meijer
       dr. M. Merkx
adviseur: dr. L.-G. Milroy
This work was co-supported by the Netherlands Organisation for Scientific Research via Gravity program 024.001.035

A catalogue record is available from the Library of the Eindhoven University of Technology
ISBN: 978-94-6108-797-3

Cover design: Inga Maria Tharun
Printed by: Gildeprint Drukkerijen, Enschede, the Netherlands
Table of contents

Chapter 1
Protein-protein interaction fine-tuning by estrogen receptor post-translational modifications

1.1 Introduction
   1.1.1 Nuclear receptor protein family
   1.1.2 The estrogen receptor
   1.1.3 Transcription factor activity of the ER
   1.1.4 Non-genomic signaling by E2 and ER
   1.1.5 ER as drug target in cancer

1.2 ER structural organization
   1.2.1 Overall domain organization
   1.2.2 Ligand binding domain
   1.2.3 Hinge region
   1.2.4 F-domain

1.3 Estrogen receptor post-translational modifications
   1.3.1 Phosphorylation of tyrosine 537 in ERα and tyrosine 488 in ERβ
   1.3.2 Modifications of the ERα hinge region
   1.3.3 ERα F-domain penultimate threonine phosphorylation

1.4 Investigation of post-translational modifications

1.5 Aim and outline of this thesis

1.6 References

Chapter 2
Ligand-independent activation of the estrogen receptor by a single PTM

2.1 Introduction

2.2 Results and discussion
   2.2.1 Protein semi-synthesis of ERα and ERβ ligand binding domains
   2.2.2 Quantification of the tyrosine phosphorylation effects on H12 dynamics and cofactor interaction
   2.2.3 On-chip cofactor-interaction screening
   2.2.4 Circular dichroism of L11-H12 peptides
   2.2.5 Solution NMR of the 15N-labled ERβ LBD

2.3 Conclusion

2.4 Experimental section

2.5 Acknowledgements

2.6 References

Chapter 3
Phosphorylation-induced interaction between the estrogen receptor and Src-kinase

3.1 Introduction

3.2 Results and discussion
   3.2.1 Expression and purification of cSrc and vSrc SH2-domains
   3.2.2 Characterization of the ER-Src kinase interaction via Fluorescence polarization
   3.2.3 Co-crystallization of vSrc SH2 domain with phosphorylated pERβ LBD
   3.2.4 Solution NMR of cSH2

3.6 References
Protein-protein interaction fine-tuning by estrogen receptor post-translational modifications

Abstract. The estrogen receptor (ER) is a nuclear hormone receptor, mediating the response of cells and tissues to estrogens which are of central interest in the development of primary and secondary reproductive organs. The ER also has regulatory functions in non-apparent gender-affected tissues such as the skeletal and vascular systems. Next to its regulation by its natural hormone ligands, the versatile ER actions are strictly modulated by post-translational modifications (PTMs). The crucial involvement of ER in vital cellular functions is reflected in ER dysregulation in several diseases, amongst which different types of cancer. About 70 % of all diagnosed breast cancers depend on functional ER for their progression, which makes ER a highly relevant drug target. Unfortunately, resistance against ER-antagonists presents a common challenge in endocrine therapy. Different pathways to endocrine resistance have emerged, often connected to an altered ER-regulation by PTMs. The detailed molecular basis of the influence of specific PTMs on ER function and structure, however, remained elusive. This is mainly due to the inaccessibility of homogeneously modified ER samples. This dissertation summarizes the explored semi-synthesis approaches for the introduction of PTMs into the ER to study their direct effects on ER structure, function and protein-protein interactions in a controlled in vitro environment by biophysical techniques.
Chapter 1

1.1 Introduction

Nuclear receptors (NRs) play a vital role in the differentiation of cells and the regulation of physiological functions. They are involved in processes like tissue differentiation during development and homeostasis in adult organisms, e.g. by regulating cell-growth and metabolism.\textsuperscript{1} The NR protein family counts 48 members in humans and is present at varying complexities throughout virtually all metazoan organisms.\textsuperscript{2,3} The substantial participation in fundamental physiological functions implies severe malfunctions upon dysregulation of NR activity, which leads to diseases such as various types of cancer, cardiovascular diseases or insulin resistance.\textsuperscript{1} Consequently, NRs belong to the most common targets in drug design, which has enriched our medicinal toolbox successfully. The most heavily targeted NR is the estrogen receptor (ER). Its small and hydrophobic natural ligand estradiol has efficiently served as scaffold for new drug design for wide variety of drugs ranging from life-style drugs to molecules applied in endocrine therapy, often successfully applied for the treatment of cancers. Unfortunately drug resistance forms a common problem in endocrine cancer therapy.\textsuperscript{4,5} Since the discovery of the two ER subtypes ER\textsubscript{α} as an estrogen-binding factor in the 1950s and ER\textsubscript{β}, four decades later, our understanding of the molecular and cellular mechanisms involved in the actions of these hormone receptors has evolved rapidly up to a point at which the various types of breast cancer start to become a prototype of individualized disease diagnose and treatment,\textsuperscript{6} with the prospect of significant improvement of the survival chance of patients. However, since the perceived complexity of a system increases upon increasing the resolution by which it is examined, many aspects of estrogen receptor function remain unanswered. In order to make rational drug design more effective and to be able to explore new approaches in NR targeting, e.g. to be able to meet the need of individualized cancer treatment or to better understand the deficiencies of existing drugs in the first place, mechanistic questions on how ER is affected by cellular signaling and direct cross-talk with other proteins needs to be explored. This dissertation aims to contribute to this process by investigating the effect of ER-modifications on the protein itself as well as on the interaction with other proteins from the ER-signaling network at molecular and atomic resolution.
1.1.1 Nuclear receptor protein family

NRs share high homology in structure and domain organization. The 48 members, that have been identified in humans for this protein family, can be divided into three groups of receptors based on the type of ligand they respond to. Next to the classical steroid hormone receptors, such as the estrogen and androgen receptors (ER and AR, respectively), a group of retinoid x receptor-dimerizing receptors has been identified, amongst which the thyroid hormone receptor. The third group refers to the orphan receptors, for which no cognate ligands have been identified (yet). This chapter will be mainly focused on the estrogen receptors (ER) and refer to the steroid hormone receptor NR subfamily. The sex hormone specific steroid receptors AR and ER, are involved in the development of primary and secondary reproductive organs but also have regulatory functions in non-apparent gender-affected tissues such as the skeletal and vascular systems. Furthermore, NR-related orphan receptors, for whom no native ligands are identified, add to the plethora of NR activities. With respect to the entire NR family, ER is by far the most heavily studied member of this physiologically and pathophysiologically important group of proteins. NRs are traditionally referred to as ligand-activated transcription factors. In the absence of their activating ligands, NR activity is silenced by various mechanisms, amongst which the association of the ligand-free NR with transcription-repressing protein complexes. Once the cellular concentration of activating ligand (agonist) rises, NRs bind their specific ligands with high affinities and undergo conformational changes which release the repressing protein interaction partners and induce their stable localization in the nucleus, where they bind to specific response elements as functional dimers. The NRs agonistic conformation allows the recruitment of coactivator proteins, which catalyze the formation of a gene transcription enhancing chromatin environment and the recruitment of the transcription machinery. The complexity of gene transcription regulation by NRs is underlined by the reported specific gene transcription repression on ‘negative’ promoter sequences by agonist-bound NR, which can in this context induce the recruitment of corepressors protein. Every individual step in this activation cascade is highly regulated by (second messenger) small molecules, protein-protein-interactions, and post-translational protein modifications (PTMs) and eventually leads to the regulation of the respective NR target gene transcription. Up- or down-regulation of target genes is followed by their specific cellular effects, such as inducing cell growth by the expression of growth factors. This heavily investigated path of ER activation reflects a rather simplified and static picture of ligand-dependent localizations and
(in-) activation in cytoplasm and nucleus as well as the historically originated focus on the ER function as transcription factor (Figure 1.2). Our understanding of ER function is, however, gradually complemented by evidence for a much more versatile and more dynamic role of ER in the cellular involvement.18,20,21 In fact, many NRs have already been identified to play a part in non-genomic cellular signaling networks.21–23

1.1.2 The estrogen receptor

The ER plays a vital role in healthy tissue development and function by mediating the physiological effects of its natural ligands such as 17β-estradiol (E2) within its target tissues. These cover for example the female and male reproductive tracts, mammary glands, the neuroendocrine system, and peripheral tissues like the skeletal and cardiovascular system.1 The importance of ER function was - next to countless other studies - illustrated by the generation of ERα and β knock-out mice.1,24,25 The knock-out of both ER-subtypes in mice allowed the development of viable mice of both sexes, yet with major deficiencies ranging from defects in proliferation, differentiation and growth factor signaling of the female reproductive tract and mammary gland to the affection of behaviours in male and female mice.1 ERα was first described as a specific E2-binding factor, identified by tissue specific accumulation of tritium labeled E2.26,27, These studies determined high E2 affinity and tissue specificity in the absorption of E2 and quantified the cellular response to the hormone by measuring the increase in RNA synthesis and cell growth.27 The importance of functional groups of steroids and the significant modulatory potential of chemically modifying E2 e.g. by fluorination was acknowledged already before the characterization of the receptor protein ERα.28–30 The resulting development of antiestrogens, which owe their name to the observation that they counteract the effects of E2 in its target tissues in a dose-dependent manner, largely contributed to the understanding of ER function at cellular, molecular and atomic resolution and increased the interest in ER as a target for the development of compounds with clinical potential.27,31–35 A major leap towards understanding the mechanism by which ER mediates the effect of its ligands was achieved following the cDNA cloning of the ER and other NRs, which revealed the conserved A/B-C-D-E-F domain organization of the NR family members (Figure 1.1).36–39 Furthermore, specific monoclonal antibodies against ERα were developed around the same time, which allowed the determination of ER cellular localization to be mainly nuclear also in the absence of E2.40,41 Since the observation of the cytosolic localization of ER in a heat-shock protein
complex, however, cellular ER localization remained an issue of controversy even though the initial observations of a predominantly nuclear localization were later on confirmed by confocal fluorescence microscopy with a GFP-tagged ER variant. This controversy implies a non-uniform division of potentially various ER-variants and tasks throughout different loci within the cell and different cell-types, which reflects their versatile roles in genomic and non-genomic signalling.

Figure 1.1 Schematic overview of the ER domain organization with a variable N-terminal domain (A/B), the highly conserved DNA binding domain (DBD, C), the variable hinge region (D), the conserved ligand binding domain (LBD), and a non-conserved F-domain (F). Sequence identity between ERα and ERβ in percent is based on the amino acid sequence of both subtypes. The underlined regions refer to cofactor interaction regions (Co-F) and regions, involved in receptor dimerization (Dimer).

The cDNA cloning also allowed for the recombinant expression of ER and other NRs, resulting in detailed molecular characterization of the receptors. These revealed that ER forms homodimers which stably bind to their DNA estrogen response elements (ERE) induced by highly conserved mechanisms of target gene regulation. Later on, heterodimerization of ERα and β turned out to be of cell-type-dependent functional importance. The functional characterization of the ER furthermore revealed the existence of two distinct activation mechanisms, either ligand-independently or in a ligand-dependent manner which were assigned to two different activation functions within the protein (AF-1 and AF-2, respectively; Figure 1.1). These distinct activation paths together with the discovery of NR coregulators, which account for cell- and promoter-specific response to ER activation, greatly broadened the scope of identified ER effects. All functional studies on ER thus far had been performed in eucaryotic cells or with their cell-extracts. Recombinant expression of full-length ER in insect cells has been reported, yet the production of functional full-length ER in bacterial cell cultures designated to yield milligram scale pure protein for
studying the receptors biophysical properties remains unachieved up until today. The isolation of individual protein domains, however, such as the receptors DNA binding domain (DBD) and the ligand binding domain (LBD) from efficient *E. coli* expression systems, provided the entrance to study the biophysical characteristics of these domains at high spatial resolution, yielding invaluable insights on the structural and functional organization of the ER and other NRs. About four decades after the first description of “estrophilin” - the estrogen binding factor which was later on termed ERα - a second ER subtype (ERβ) was identified. Both ER subtypes are individual proteins, encoded by different genes, which are located on different chromosomes. Nowadays, several less studied variants originating from alternative splicing products of these two genes are known, as well as a class of ER related receptors (ERRs), adding another level of complexity to estrogen signaling. Subsequent extensive studies of both ER subtypes revealed that the roles of ERα and β are non-redundant. Both subtypes have individual tissue specific expression patterns and co-regulate a number of tissues amongst which mammary gland, thyroid and brain tissue. They exert subtype specific effects as transcription factors mediated by different coactivator interactions and the association with distinct EREs as well as their different responses to ligands. In female breast tissue for example, ERα activity appears to be vital for mammary development, which is not affected by ERβ-knock out. ERβ has also been reported to antagonize ERα-regulated E2-induced effects, like for example the proliferation in mammary gland development. The role of the fine balance between ERα and ERβ activities has consequently been studied in cancer and its disturbance has been linked to the clinical outcome of cancer therapy. In fact, the expression ratio between ERα and ERβ in breast cancer patients could be directly correlated to the outcome of endocrine therapy. The origin of differential effects between the two ER subtypes can in part be accounted directly to the receptor inherent differences and in part to the indirect effect of their different signalling partners. The sequence similarity between ERα and ERβ is the highest in their DNA binding domain (*Figure 1.1*). The ligand binding domain is structurally conserved, yet shares only moderate sequence similarity between ERα and β, resulting in about a 20 % smaller ligand binding pocket for ERβ, logically leading to differential ligand effects between the two subtypes. The large variations in the regulatory N-, hinge-, and F-domain significantly add to the receptor-inherent differences. Next to these sequence and structure induced modulatory characteristics, differential regulation by
subtype specific expression patterns, DNA-ERE-specificity, protein protein interactions and PTMs contributes to the different functional involvement of the two ER subtypes.\textsuperscript{1,51,78,95–97}

1.1.3 Transcription factor activity of the ER

ER transcription factor activity can be sub-divided in three different pathways of action (\textbf{Figure 1.2}).\textsuperscript{98} These different ER transcription factor pathways are categorized as follows: (1) Classical direct activation of ER by its natural agonists, resulting in target gene transcription (\textit{vide infra}). (2) Growth factor signaling and its downstream mitogen-activated protein kinase (MAPK) pathway induce ER transcription factor activity in an ER-ligand-independent manner via direct N-terminal phosphorylation of ER.\textsuperscript{99,100} (3) Estrogen response element (ERE) – independent action of ligand-bound ER by recruiting cofactors to alternative promoters via direct interactions with other transcription factors, amongst which orphan receptors but also unrelated transcription factors.\textsuperscript{101–103} This latter mode of ER transcription factor activity has been estimated to account for \textasciitilde30 % of all ER-regulated genes.\textsuperscript{101–103} In its classical role of ligand-activated transcription factor (\textbf{Figure 1.2}, (1)), ER regulates target gene transcription in a highly dynamic manner. ER initiates the cellular response to the hormone signal by mediating an adapted target gene expression pattern.\textsuperscript{92} As a mechanism of transcriptional control, stochastic binding events of ER to its estrogen response elements (EREs) located in target gene promoters have been proposed, by which ER cycles between DNA-association and dissociation, allowing continuous surveillance of the potential changes in hormone levels in the cell.\textsuperscript{3,18} Binding of agonistic ligand and receptor dimerization on its ERE allows coactivator recruitment and subsequent target gene transcription. Many NR coactivators are known to date,\textsuperscript{60,104,105} often equipped with chromatin remodeling catalytic activity, such as histone acetyl-transferase- and histone methyl-transferase- activities, which are subsequently actively involved in transcription activation via transcription enhancing histone-modifications.\textsuperscript{106,107} In the absence of agonist or when bound to (partial) antagonists, transcription repression is maintained by the direct interaction of ER with corepressors, such as the NCoR or SMRT, which recruit histone-deacetylases (HDAC) and thereby induce a repressive chromatin environment.\textsuperscript{107} In contrast to coactivator proteins, which often possess a modular domain structure, NCoR, SMRT, and other NR corepressors, like RIP140,\textsuperscript{108,109} are predicted to be mostly intrinsically unstructured yet harboring a histone- and a HDAC-binding region.
Figure 1.2 Schematic overview of the ER transcription factor activities, directly activated by classical agonists (E2) and resulting in target gene regulation via estrogen response elements (ERE) (1) or indirectly via interactions with other transcription factors (TF) and their cognate promoter (TF-E) (2). (3) Activation of ER by phosphorylation (P) via kinase signaling cascades, e.g. triggered by growth factors (GF) and their receptors (GF-R).

A variety of proteins have been described which directly enhance or repress NR-regulated gene transcription and thereby qualify for the description of NR coregulators. Most of these proteins also coregulate gene transcription, which is controlled by other transcription factors than NRs. The majority of these diverse proteins are built up from multiple domains, which determine their overall functions. They exert as versatile functions as chromatin remodeling, chromatin binding, NR or other transcription factor linking, and catalyzing activating or inhibiting PTMs of transcription factors or other coregulators. Approximately half of the diverse group of NR coactivator proteins carries a sequence motif, which is generally referred to as ‘NR box’ (LXXLL or FXXLF, in which X is any amino
The definition of NR coactivators often specifically refers to this subgroup of proteins. The latter nomenclature will also be used throughout the following chapters, however keeping in mind that the discussed regulatory mechanisms refer to a subgroup of coregulators. The ligand-regulated interaction between NR coregulators and their NR-interaction partners has been shown to occur via direct interactions with the receptors LBD and does not require any of the other NR domains (Figure 1.3), yet the N-terminal AF-1 of ER\textsubscript{α} has also been shown to harbor coregulator binding regions. Variations in the sequence embedding the ‘NR box’ binding motif, influence NR-coregulator interaction affinity and kinetics, which is proposed to have regulatory function. Furthermore, ER subtype specificity in coregulator interaction has been observed in many cases, contributing to the differential effects of the two ER subtypes and illustrating their often unlike response to the same ligands. The formation of a coregulator binding surface in a ligand-modulated manner by conformational rearrangements in the LBD allows the specific protein-protein interaction via hydrophobic and polar interactions (vide infra). Due to the prospect of downstream gene transcription and consequently e.g. tumor growth inhibition by newly developed ligands with potential clinical relevance, the quantification of the NR-coregulator interaction has become a tool to screen the inhibitory potential of NR-ligands. Just like ER-activity, cofactor-activity is regulated by post-translational modification through signal transduction-involved modifying enzymes. This direct effect of non-genomic signalling cascades on ER cofactor activity, which is consequently affecting the ER-target gene transcription has been seen as one way of convergence between ER and e.g. growth factor induced non-genomic cellular signalling. Due to the large diversity of ER-involvement in non-genomic signalling cascades, mechanisms of the various roles and regulation of ER in non-genomic networks remains to be explored.

1.1.4 Non-genomic signaling by E\textsubscript{2} and ER

Next to the classical mechanism of target gene regulation by ER as ligand-activated transcription factor, its involvement in rapid and often referred to as non-genomic signaling becomes of increasing interest in order to gain full insight on the various roles of ER activity in cellular regulatory processes. Since the ultimate effects of these rapid signaling cascades often still effect the regulation of ER target gene transcription, the terminology ‘rapid signaling’ represents a more adequate naming. The rapid effects of E\textsubscript{2} and ER were initially
identified based on the observation that some responses within the target tissues of E2 occur rapidly -within seconds,\textsuperscript{120–123} whereas others need a longer time frame of hours to establish.\textsuperscript{76} The rapid responses to E2 are versatile and it has been shown that a variety of receptors is involved in their mediation - it cannot even be excluded that ER-independent E2-initiation of signalling cascades is in part responsible for some reported ligand-assigned effects.\textsuperscript{76} However, many cell type and cell-state-specific rapid actions of E2 have been identified, which are associated with the direct involvement of ERs. Examples are the E2-induced immediate increase of second messengers inositol triphosphate (IP\textsubscript{3}) and calcium, enabling for example ER-transactivation at low E2 levels.\textsuperscript{120,124–126} Fast effects of E2 can be mediated by membrane-localized ERs.\textsuperscript{120} These have been shown to originate from the same ER transcripts as nuclear ERs and are localized to the cell membrane by palmitoylation.\textsuperscript{127–130} Absence of the membrane-localized ER has recently been shown to result in severe abnormalities in organ development and are of comparable significance for proper tissue and organ development as the entire ER population,\textsuperscript{1,131} underlining the importance to broaden the detailed understanding of non-classical ER-functions.

Direct ER membrane localization and ER-ligand binding are, however, not necessarily required for ER involvement in fast cellular signalling events. In fact, even though the general awareness of ligand-independent involvement of ER in signaling cascades appears to be lower than that of ligand-bound ERs, evolutionary models, suggesting a late evolution of ER ligand-control, support the high probability of a versatile ligand-independent involvement of ER in signaling cascades.\textsuperscript{132,133} An example for estrogen-independent involvement of ER in a cytosolic signaling cascade is the reported complex formation between ER and cSrc kinase in different contexts. In mammary cancer cells, ER\textsubscript{α} has been shown to interact with cSrc kinase via its phospho-tyrosine binding SH2 domain in a phosphorylation-dependent manner and the same was shown for ER\textsubscript{β} in prostate cancer cells.\textsuperscript{134–137} The formation of a ternary complex between ER, cSrc kinase and the androgen receptor has been connected to enhanced cell proliferation, via the activation of the Src/Raf-1/Erks-pathway.\textsuperscript{135} Clinical relevance of this interaction became apparent and the inhibition of the latter antagonists for either ER or AR has been reported.\textsuperscript{135} Yet, cSrc kinase has been reported to be over expressed in several types of cancers amongst which types of ER-positive breast cancer and has been shown to stimulate cancer progressing processes like cell proliferation, metastasis and angiogenesis.\textsuperscript{138–141} In this context, E2-bound ER\textsubscript{α} has an activating effect on cSrc kinase and both proteins individually have consequently been
identified as drug targets in the treatment of these specific ER-positive breast cancer cases, resulting in the development of Src inhibitors. A combination of ER antagonist with cSrc kinase inhibition has shown to be more promising than the individual targeting of one of the two proteins. Also the direct or indirect inhibition of cSrc kinase-activation by ER can be considered as a valuable inhibitory path for cancer treatment. The molecular determinants of the ER-cSrc SH2 interaction are explored in molecular and atomic detail in chapter 3.

1.1.5 ER as drug target in cancer

Many parameters, such as family history and age in combination with exposure times to estrogenic compounds, affect the risk of an individual to develop breast cancer. Hormone responsive breast cancers, specifically ER-positive breast cancers, have been treated by interfering with the endocrine system successfully even before antiestrogens or agents interfering with physiological estrogen-synthesis were available by removing the estrogen producing ovaries and thereby reducing the estrogen levels. Since estradiol is not only produced in the female reproductive organs but also in for example breast tissue, systematic inhibition of estradiol-synthesis catalyzing aromatases represents the modern equivalent to the effects of organ-removal. Together with aromatase inhibition, ER-inhibition by anti-estrogenic small molecules - such as tamoxifen - are two most widely applied therapies of estrogen responsive breast cancers. Since 70% of all breast cancers are ER-positive and thus can in principle be treated by interference of that activity, the research and development of improved and more specialized drugs stays of interest. Its relevance is highlighted by the fact that about 30% of the diagnosed breast-cancer patients are resistant to anti-estrogen therapy and about 40% of the treated patients acquire resistance throughout the endocrine treatment. Many mechanisms and cell characteristics behind innate and acquired tamoxifen resistance for example have been described and analyzed for their potential as biomarkers for the prediction of the best choice of treatment. One extensively studied example for such a biomarker is the phosphorylation status of serine 305 in ERα. Phosphorylation at this position has been shown to be stimulating for ER-activity and renders the partial antagonist tamoxifen from an ER-antagonizing inhibitor into an agonistic ER-activator. Furthermore, hyperphosphorylation of S305 by protein kinase A (PKA) in breast cancer tissues has been observed, resulting in an ER with elevated activity and adverse regulation by tamoxifen. Molecular mechanisms underlying this specific pathway to tamoxifen resistance appear to be
mediated by phosphorylation-induced conformational changes in the ER, resulting in altered orientations towards coactivators.\textsuperscript{159,160} Tamoxifen agonism has turned out to be a subtype specific ligand-effect, mediated by ER\textsubscript{α} AF-1 and hinge region,\textsuperscript{94,160} but not by ER\textsubscript{β} due to the lacking ability of ER\textsubscript{β} to interact with coactivators in the tamoxifen-bound state.\textsuperscript{94} These observations undeline the significance of understanding the direct impact of individual PTMs on protein structure and function in order to understand drug resistance and by that contribute to the efficient rational design of new drug molecules.

Several alternative concepts to overcome or circumvent resistance to endocrine therapy have been reported. These approaches include for example agents, which interfere with the ER\textsubscript{α}-cofactor or ER\textsubscript{α}-DNA interaction and thus inhibit ER\textsubscript{α} function and are referred to as nuclear receptor alternate-site modulators (NRAMs).\textsuperscript{161} The context-dependent effects of SERMs (selective estrogen receptor modulators), like tamoxifen and raloxifen,\textsuperscript{77} have been intended to be circumvented by the development of ER-subtype specific molecules.\textsuperscript{162–166} Also the indirect effects of ER-signaling via ERE-independent stimulation of other transcription factors have been targeted for drug design already.\textsuperscript{83} SERDs (selective ER\textsubscript{α} down-regulators) are aimed to trigger ER\textsubscript{α} degradation and thereby inhibit its signaling activity.\textsuperscript{167,168} A recently developed concept for the development of new ER-anagonizing molecules aims for the stabilization of protein complexes which have an inhibitory character on ER-activity.\textsuperscript{169,170} The phosphorylation-dependent interaction between ER\textsubscript{α} and 14-3-3 proteins is explore in chapter 5.
1.2 ER structural organization

1.2.1 Overall domain organization

As mentioned earlier (Figure 1.1) the estrogen receptors share the highly conserved NR domain organization.\(^{36-39}\) With the N-terminal domain, being involved in ligand-independent transactivation, the DBD,\(^5^3\) establishing direct interactions with the ERE-elements in ER-controlled gene promoters, the hinge region, concerting the activities of AF-1 and AF-2,\(^9^4\) the LBD,\(^1^7^1\) harboring the ligand-regulated AF-2 next to a dimerization interface for homo- and hetero-dimerization,\(^9^0\) and the F-domain,\(^4^6\) being involved in ligand-effect mediation. Further on, the focus of this introduction will be on the domains, which make direct part of this dissertation.

1.2.2 Ligand binding domain

The LBD represents the most intensively studied domain of the ER. This is most-certainly explained by the three central functions of the LBD, covering the ligand binding pocket, a cofactor interaction surface and a dimerization interface. The latter facilitates the direct interaction with other ER LBDs and thereby facilitates the dimeric binding mode of the ER to its DNA-recognition motif.\(^3^7^3^5^3^4\)

Figure 1.3 Crystallographic protein domain structures of different LBDs with functionally important H12 in green, coactivator peptide in blue and ligand in black (sticks). A: Dimer of the ERβ LBD (monomers in light and dark gray) in complex with the ligand E2 and cofactor peptide (PDB ID: 3OLL). B: RXRα LBD monomer in the ligand-free state (PDB ID: 1LBD). C: ERα LBD monomer in complex with OHT (PDB ID: 3ERT)
As its name implies, the LBD furthermore accommodates a ligand binding pocket, to which natural high-affinity ligands bind. The activating or inhibitory effects of individual ligands are directly translated by the induction of receptor structural rearrangements. The most C-terminal helix of the LBD (H12) has been shown to play a crucial regulatory role in the transmission of ligand-effects. In fact, H12 orientation fine-tuning is known to directly affect the interaction with coregulator proteins, since the agonistic orientation of H12 has been shown to complement the coregulator interaction surface (AF-2), representing the third central function of the ER LBD. Many crystal structures have been reported of different steroid hormone receptors (SHRs), revealing a conserved LBD fold and the influence of agonist and antagonist binding on the protein structure.\(^{31,35,91,172,173}\) This conserved domain fold consists out of two short anti-parallel \(\beta\)-strands and 12 \(\alpha\)-helices (Figure 1.3).\(^{90,171,174}\)

The importance of the most C-terminal H12 has become clear from biophysical, cellular and many crystallography studies of SHR LBDs in the presence of different ligands and cofactor peptides as well the structural data, obtained from ligand-free LBDs, which has only been obtained for a few LBDs so far (Figure 1.3 B).\(^{16,17,20,175,176}\) In the latter, H12 protrudes from the core LBD-fold, supporting the assumption that H12 folds like a lid over the ligand binding pocket in the presence of ligand, yet resides in an open conformation in the absence of ligand.\(^{16,31,35,177}\) The precise positioning of H12, depending on the ligand, has been correlated to ER transactivation or inhibition. Upon binding of an agonist, H12 relocation complements the coactivator binding groove, which is consisting of residues from helix 3 and helix 4 as well the charged residues K362/314 and E542/493 in ER\(\alpha/\beta\),\(^{178,179}\) forming a dipolar ‘charge clamp’,\(^{16,178,180,181}\) which can undergo hydrogen-bonding interactions with the dipolar coactivator helical binding motif, surrounding its NR box binding motif ‘LXXLL’.\(^{110,113,115,182}\) The glutamic acid residue, which is making part of the charge clamp is located in H12, resulting in a formation of the charge clamp only in the closed agonistic conformation of H12. In the context of the packing of a protein crystal, the antagonistic tamoxifen induces an alternative orientation of the H12, in which the charge clamp is not formed and H12 is blocking the interaction surface for coactivator binding.\(^{31,35,177}\) General alterations in H12 dynamics by ligands has also been demonstrated by in-solution fluorescence assays.\(^{183–185}\) The potential effects of ER phosphorylation in the C-terminal LBD, precisely at the N-terminal loop 11-H12-transition, on cofactor interaction and H12-dynamics is explored in chapter 2.
1.2.3 Hinge region

The ERα hinge region is assumed to be an overall unstructured region of the receptor protein, which acts as a platform for interaction with modifying enzymes as well as other activity regulating proteins and contains a nuclear localization sequence. Its location between the DBD and LBD as well as its length and sequence have been demonstrated to be of great importance for the concertation of the N- and C-terminal activation functions of ERα. ERβ lacks a functional AF-1 and its hinge region has a fundamentally different character, which is by far less well understood, than that of ERα hinge region. In addition to the recently characterized crucial role in AF-concertation, ERα hinge region PTMs have been implied to be involved in ERα transactivation regulation (vide infra). The underlying mechanisms of the plethora of hinge region PTM-effects however is poorly understood to date, with the exception of clinically highly relevant S305-phosphorylation. The effects on several hinge region modifications on hinge region structural character and an approach towards introduction of hinge region modifications into ERα protein constructs are explored in chapter 4.

1.2.4 F-domain

Not much is known about the NR F-domains to date. There is hardly any structural information available and its biological function and biophysical character has not been very intensively studied in general. The F-domains of different NRs are highly variable. Also the F-domains of ERα and ERβ differ significantly in length and sequence identity. Its location at the very C-terminus of the ER LBD and thus immediately following the regulatory H12 implies a role of the F-domain in fine-tuning of ligand-responses. This has been demonstrated by several studies. Furthermore, the F-domain has been associated with an inhibitory effect on ER-dimerization, of which at least one underlying mechanism has recently been demonstrated to be mediated by phosphorylation-dependent 14-3-3 interaction. Chapter 5 covers the introduction of the F-domain threonine-phosphorylation into an ERα protein construct and the biophysical in vitro characterization of the interaction between 14-3-3 and ERα F-domain in perspective of this interaction providing a promising drug target protein-interface.
1.3 Estrogen receptor post-translational modifications

PTMs are chemical modifications of cellular biomolecules and represent specific temporal signals for (intra-)cellular communication, which provide a highly utilized tool for cells and tissues in fast, often reversible and optional long-term marking of (sub-)cellular network members. PTMs regulate virtually any step in complex cellular systems ranging from the turnover of network members (ubiquitylation), their transcription levels (histone modifications) to the fine-tuning of one-another’s activities via direct protein modifications such as phosphorylation, acetylation or methylation.\textsuperscript{197–199} Also the ER is highly regulated by PTMs (Figure 1.4).\textsuperscript{96,200–208} Even though less is known about ERβ modifications, PTMs are proposed to present one of the mechanisms by which subtype-specific receptor regulation is mediated.\textsuperscript{97} Furthermore the phosphorylation status of different ER residues amongst which S305 (chapter 4) and Y537/Y488 (in ERα/ERβ, chapter 2 and 3) have been reported to be promising biomarkers for the prediction of cancer treatment success as well as the determination of the levels of their modifying enzymes.\textsuperscript{4,139,141,154,156,158,209} This section summarizes the available information on the relevant ER modifications.

\textbf{Figure 1.4} Schematic overview of ERα post-translational modifications. Generally activating modifications are indicated by empty shapes and overall inhibiting modifications are indicated by filled shapes. The different types of amino acid side chain modifications are symbolized by different shapes, whit phosphorylation (oval), methylation (triangle), acetylation (rectangle), sumoylation (asterisk), ubiquitylation (U), palmitoylation (underlined) and the glycosylation (italic). This overview is adapted from literature\textsuperscript{96} and has been complemented.

1.3.1 Phosphorylation of tyrosine 537 in ERα and tyrosine 488 in ERβ

Tyrosine 537 (Y537) in ERα and the analogous tyrosine 488 (Y488) in ERβ are phosphorylated in a ligand-independent manner by Src-kinase.\textsuperscript{97,135,210–212} in a context-specific manner.\textsuperscript{204} Both tyrosines are located immediately at the loop-helix transition - or “capping motif”\textsuperscript{213} - N-terminal of H12.\textsuperscript{31,214} Due to their location in a highly involved
Protein-protein interaction fine-tuning by estrogen receptor post-translational modifications

protein region in receptor transactivation regulation, phosphorylation of Y537 and Y488 are suspected to exert an important contribution to the dynamics of the C-terminal H12 and thereby have regulatory influence on ER activity. When it was first discovered, the phosphorylation has also been reported to enhance receptor dimerization and thereby increase ER-transactivation. A series of mutations of the tyrosine residues of interest or mutations of amino acid residues in the immediate vicinity have been shown to have dramatic effects on receptor transactivation, emphasizing the highly specific function of this protein region. The actual phosphorylation status of the tyrosine of interest has been implied to be of functional importance for the ER with respect to ligand binding affinity and kinetics, DNA binding, transactivation and PPIs. Mutagenesis studies of the Y537 to serine, resulted in an up to 90 % consecutive activity of the receptor in the absence of E2, which is most probably mediated by its enhanced agonist-like conformational stability and increased co-activator interaction. A small but significant tendency was also reported for the Y537A, while the introduction of a bulky amino acid residue (phenylalanine) instead of the potentially phosphorylated tyrosine side chain, was reported to have destabilizing effects on the receptor, increased the \( k_{\text{off}} \) rate of the E2 binding kinetics and furthermore abolished the stimulating effect, which was observed for other mutations. Mutation of the tyrosine of interest to an aspartic and glutamic acid, supposedly mimicking the side chain properties of phosphorylated tyrosine, have shown a significantly increased activity of the ER\( \alpha \) and \( \beta \) mutants in the absence of ligand and even render the receptor consecutively active, which can be antagonized for by anti-estrogens for both ER-subtypes. Intriguingly, a number of clinical studies have recently revealed ER\( \alpha \) Y537 mutations, amongst which the earlier characterized Y537S mutant, to be activating in hormone-resistant metastatic breast cancer. The reported occurrence of Y537 mutations is indicating a hot spot of mutations at position Y537 in metastatic breast cancer patients, which have been treated with endocrine therapy for longer than 3 months and have been associated with the relapse of the disease, implying a mechanism for acquired resistance to endocrine therapy.

The important insights sketched above have mainly been obtained by cell-based and clinical studies. The exact molecular mechanism which underlies the final cellular effects has remained unresolved. Several studies have indicated that H12 dynamics and structural integrity underlie a fine-balanced regulation between active and inactive states. Molecular elucidation of the effect of tyrosine phosphorylation on
H12 dynamics regulation and coactivator binding is therefore of great interest, regarding the reported clinically relevant mutations, overexpression of tyrosine kinases responsible for ER phosphorylation in cancer cells, and new ligand development. Both phosphorylated ER subtypes have furthermore been reported in cellular assays to interact directly with the cSrc tyrosine kinase in a phosphorylation-dependent manner. The interaction of these two proteins has however not been characterized and quantified in vitro before. A major limitation for performing a molecular analysis of the effects of tyrosine phosphorylation on ER structure and function as well as its effect on protein-interaction has been the inaccessibility of well-defined phosphorylated ER constructs for biochemical and structural studies. In chapter 2, a semi-synthetic approach on the introduction of the tyrosine phosphorylation of interest into ERα and ERβ is therefore explored, resulting in a library of ER protein constructs which have been used to investigate a series of intra- and intermolecular effects of the modification of ER behavior as well as for the quantification of the ER-cSrc tyrosine kinase interaction (chapter 3).

1.3.2 Modifications of the ERα hinge region

The ERα hinge region is the most densely post-translational modified region of the receptor protein (Figure 1.4), implying that one of its major functions is the fine-tuning of receptor activity and signaling network-involvement via these modifications. ERβ hinge region differs greatly from ERα in function and structure and has until now not been reported to be this highly modified. The ERα hinge region modifications however provide an excellent example of the plethora of effects, which can be achieved with chemically identical modifications. The N-terminal end of the hinge region, situated immediately C-terminal of the ERα DBD harbors lysine-acetylation sites, which enhance DNA-binding upon acetylation whereas the acetylation of lysines 299, 302 and 303, located at the C-terminal end of the ERα hinge region, reduces ER-transactivation. In fact, all mentioned lysines are acetylated by the same acetyl-transferase (p300), emphasizing the complexity of PTM-regulation. The work, described in chapter 4 focuses on the interplay and individual effects of C-terminal hinge region modifications K299, K302, and K303 as well as the activating S305-phosphorylation on hinge region structural character. The opposing effects of acetylation and phosphorylation in such close sequence vicinity as well as considerable regulatory cross-talk between the modifications and the involvement of S305 in endocrine resistance in cancer therapy, makes these modifications of great interest. The three mentioned lysine
residues have furthermore been reported to be regulated by sumoylation, which enhances receptor transcriptional activity. Next to enhancing or reducing ERα transcriptional activity, K302 and K303 are involved in the receptor turn-over, by being targets for ubiquitinylation. K302-methylation represents a directly stimulating modification for ERα transcription activity and cross-influence to some extent between K303 modification with K302-methylation has been reported.

S305 is located at the border between highly structured C-terminal LBD and less structured N-terminal hinge region (Figure 1.4) and its phosphorylation by PKA has been reported to enhance receptor activity. This phosphorylation-mediated up-regulation of ERα activity has been reported to often turn into a challenging problem in the treatment of ER-positive breast cancer patients, which are insensitive to endocrine therapy by either aromatase inhibition as well as treatment with ER-inhibiting antagonists like tamoxifen. The phosphorylation status of S305 as well as expression levels of PKA have therefore been investigated as a biomarker for the prediction of cancer treatment success. On a molecular level, the effect of S305-phosphorylation appears to be mediated by altered structural arrangements within the receptor domains and with respect to coactivator binding. Investigating the direct effects of hinge region modifications and their potential mechanistic implications by a controlled in vitro chemical biology approach is worked towards in chapter 4.

1.3.3 ERα F-domain penultimate threonine phosphorylation

The F-domain of the ER represents another domain, reported to facilitate ER subtype specific effects. It is located immediately C-terminal of the functionally important H12 and is thus likely to affect H12-behavior. The low extent of conservation between the F-domain of ERα and β also suggests a role of the F-domain in the mediation of ER subtype-specificity. In contrast to the hinge region or the N-terminal domain, however, the F-domain does not appear to exert its regulatory effects through a plethora of PTMs. This assumption could be proven wrong as research advances, but to date only two types of F-domain modifications have been reported. A cellular sub-population of ER has been reported to be glycosylated in a context specific manner. Another modification of the ERα F-domain - the phosphorylation of the penultimate threonine - has only recently been identified and appears to be significantly involved in ERα activity regulation. The enzyme which catalyzes T595-phosphorylation has not been identified yet, however the functional
importance of the threonine phosphorylation has been accounted for based on the phosphorylation-dependent interaction with 14-3-3 proteins and its connection to the earlier reported F-domain influence on receptor dimerization.\textsuperscript{46,169,192,195,196} The phosphorylation-dependent interaction of ER\(\alpha\) F-domain with 14-3-3 proteins is explored in chapter 5.
1.4 Investigation of post-translational modifications

Cellular methods for the investigation of the existence, origin and effects of individual native PTMs often represents the basis for the characterization of PTMs. For this purpose, cellular signaling pathways, resulting in the modification of target protein are triggered to study the PTM-effects in vivo.\textsuperscript{158} Also mutagenesis of the residue of interest in the target protein is a widely applied tool for the identification PTM-effects on protein function.\textsuperscript{229} Such a mutation of individual residues is often done with the intention of electrostatic-remodeling of amino acid residues by for example introducing an opposite charge or an amino acid, mimicking the electrostatic and physical properties of the respective modification. The role of protein modification in protein interactions can also be evaluated by the use of specific disruptors of the protein-interaction of interest by for example peptide analogues of the modified protein sequence.\textsuperscript{245,246} In recent years the identification of protein modifications by mass-spectrometry techniques has furthermore allowed systematic mapping of protein modifications and of their complex cross-talk.\textsuperscript{197,247} Next to the valuable information that can be gained from cellular systems, one major disadvantage is the complex nature of the applied system. No certainty about the assignment of the observed functions of modifications to individual specific modifications can be achieved. This is due to the unavoidable side effects of induced stimuli on the investigated system, such as parallel signaling cascades, e.g. resulting in additional target modifications. \textit{In vitro} enzymatic modification of proteins to yield the introduction of specific PTMs into recombinantly purified proteins for the investigation of individual PTMs in a controlled \textit{in vitro} environment allows for complementing information to the cellular data.\textsuperscript{210} However, this approach knows several drawbacks such as limited availability of modifying enzymes or even the lack of knowledge on the identity of the native modifier. The enzymatic introduction of specific and individual PTMs furthermore requires a single enzymatic target site within the target protein sequence in order to be site-directed. To circumvent multiple modifications of the target protein in the case of multiple enzyme recognition sequences within the target protein sequence, point mutations may thus be required in order to mask the alternative recognition sites, resulting in a non-native and potentially significantly altered target protein. Additionally, enzymatic reactions will often result in incomplete modification, due to reversibility of the catalyzed reaction or other characteristics, such as slow reaction kinetics combined with low stability of the enzyme, and thus yield inhomogeneous protein samples, which is suboptimal for the read-out of many biophysical methods, relying on in-bulk
analysis. The homogeneous introduction of PTMs into recombinantly expressed proteins can however be achieved by chemical biology approaches. The term chemical biology is often used for concepts that make use of chemical methods, such as specific reactions on biomolecules with synthetically derived components, for the investigation of a biological process or system.\textsuperscript{248–251} Protein modification by the introduction of unnatural amino acids has been one of the major achievements in chemical biology so far and has contributed to the investigation of protein PTMs.\textsuperscript{252–256} The introduction of amino acid modifications for the \textit{in vitro} study of PTMs by biophysical techniques, however, requires milligram yields of modified protein, which is untypical for the ribosomal introduction of unnatural amino acids. For the efficient, stable and high-yield introduction of PTMs into proteins, the utilization of the concept of expressed protein ligation (EPL) has proven to be a valuable protein-engineering tool.\textsuperscript{252,257–262} The EPL approach requires the production of an N-terminal peptide or protein moiety, carrying a thioester at its C-terminus and a C-terminal protein or peptide moiety, carrying an N-terminal cysteine (\textbf{Figure 1.5}).\textsuperscript{250,252} These two components can subsequently react in aqueous solutions and in the presence of a thiol-catalyst under reducing conditions and at physiological pH values to yield a native peptide bond via a chemoselective native chemical ligation (NCL) reaction (\textbf{Figure 1.5}).\textsuperscript{263}

\textbf{Figure 1.5} Reaction mechanism of the NCL reaction between a (poly-)peptide thioester and a (poly-)peptide with an N-terminal cysteine.\textsuperscript{257,263}
When combining a recombinantly expressed protein-moiety with a synthetically produced peptide sequence, the possibility for the simultaneous introduction of several diverse PTMs as well as read-out labels for the investigation of the resulting semi-synthetic protein construct by biophysical techniques, can efficiently be accomplished by EPL.\textsuperscript{259,264,265} The accessibility of protein-thioesters came along with the recombinant expression of engineered self-splicing intein-fusion protein constructs, which can yield proteins with N-terminal cysteines.\textsuperscript{266–271} Alternative methods for the production of protein thioesters by a sortase reaction as well as the introduction of N-terminal cysteines into protein-moieties have been developed, including the site-specific cleavage by proteases immediately N-terminal of the cysteine of interest.\textsuperscript{272–275} The applicability of the NCL reaction for the semi-synthetic production of post-translationally modified ER constructs will be explored in the following chapters.
1.5 Aim and outline of this thesis

Much is known about the function and structure of the two ER subtypes, in fact they represent two of the most heavily studied NRs. This results from the clinical relevance and pharmacological potential of these hormone receptors. Increasingly detailed evidence for the ER-regulation in its genomic function as transcription factor as well as a slowly growing non-genomic component of ER-function is emerging. In all these diverse facets of ER-function, the importance of ER-regulation by PTMs becomes increasingly evident. ER PTMs take care of the essential context-specific fine-tuning of ER-function by directly modulating its structural and dynamic character and by facilitating or abolishing protein-protein interactions with other proteins from the ER genomic and non-genomic signaling. The molecular basis of these PTM-effects however has not been explored in great detail to date which is due to the inaccessibility of homogenously modified protein samples. Homogeneously modified protein samples carry the enormous potential to specifically assign structural and functional effects of amino acid side chain modifications within the protein context, which will contribute to the overall understanding of the complex mechanisms of ER-activity regulation. Next to feeding the pool of knowledge, which is certainly of use for the efficient development of new effective drug molecules or for the rational prediction of the individualized benefits of existing therapies, understanding fundamental effects of amino acid side chain modifications might allow to establish basic principles, which can help interpreting entirely different systems at given analogy. The applied chemical biology approach, employing native chemical ligation facilitated introduction of protein PTMs has been proven to be very effective for the simultaneous introduction of read-out labels for biophysical protein characterization to yield detailed insight to the range of influence of a single PTM (chapter 2) and to characterize modification-dependent protein-protein interactions of the receptor which have not been quantified in vitro before (chapter 3). Furthermore, a combination of protein peptide analogues and protein semi-synthesis allows the elucidation of important, yet little characterized protein interactions and reveals the powerful influence of PTMs on intrinsic protein region structural characteristics (chapter 4 and 5). The thorough analysis of challenges, intrinsic to specific ER semi-synthesis approaches provide a base towards the introduction of highly relevant ER modification into the ER protein context (chapter 4 and 5). Altogether the described chemical biology approaches of ER PTM-effects on protein structure, function and protein-protein interaction also provide an example for the future introduction of PTMs into other NRs.
1.6 References

(7) Kumar, R.; Thompson, E. B. Steroids 1999, 64, 310–319.
(36) Kumar, V.; Chambon, P. Cell 1988, 55, 145–156.
(37) Kumar, V.; Green, S.; Stack, G.; Berry, M.; Jin, J. R.; Chambon, P. Cell 1987, 51, 941–951.
Protein-protein interaction fine-tuning by estrogen receptor post-translational modifications

Protein-protein interaction fine-tuning by estrogen receptor post-translational modifications


Chapter 1


Protein-protein interaction fine-tuning by estrogen receptor post-translational modifications


Chapter 1


32
2 Ligand-independent activation of the estrogen receptor by a single PTM

Abstract. The role of a plethora of post-translational modifications in the regulation of nuclear receptor activity has been acknowledged mainly in cellular and clinical studies. This chapter covers the dissection of ER phosphorylation and ligand-binding effects on a molecular level by making use of a semi-synthetic approach for the homogenous introduction of tyrosine phosphorylation to ERα and ERβ LBD. Simultaneously, the introduction of site-specific fluorescent or 15N-labels for the investigation of tyrosine phosphorylation effects on H12 dynamics is described. Circular dichroism spectroscopy was employed to complement the characterization of tyrosine phosphorylation impact on ER structure and dynamics. Furthermore the effect of tyrosine phosphorylation on cofactor interaction was quantified. The described data reveals a subtype specific regulation of ERα and ERβ activity. ERβ tyrosine phosphorylation induces coactivator binding affinities of the same order of magnitude as the agonistic ligand E2, whereas ERα specifically binds to corepressor peptides upon phosphorylation. These effects appear to be mediated by phosphorylation induced local disruption of H11-loop-H12 secondary structural character. The results summarized in this chapter emphasize the potential of protein semi-synthesis for the elucidation of the effect of protein modification and - given their clinical relevance - provides a base for the development of molecules, targeting the phosphorylation-regulated ER specifically.

Part of this work has been submitted for publication: Tharun, I.; Nieto, L.; Haase, C.; Scheepstra, M.; Balk, M.; Möcklinghoff, S.; Dames, S. A.; Brunsveld, L.
2.1 Introduction

Much focus has been drawn to the structure and function of the ER LBD for the design of new ligands as well as to understand the ligand-induced activation mechanism of the receptors in more detail. Numerous crystal structures in the presence of different ligands and coregulator-derived peptides have revealed a conserved α-helical rich fold of the LBD and the structural impacts of agonistic and antagonistic ligands.\textsuperscript{1-5} Together with \textit{in vivo} and \textit{in vitro} studies of NR function,\textsuperscript{6,7} a widely accepted activation mechanism of NR transcription factor-activity has been established.\textsuperscript{8} A central role in the translation of ligand-binding in the activation of the receptor is assigned to the structural rearrangement of the most C-terminal α-helix (H12), of the ER LBD. Upon binding of agonists, such as E2, the spatial orientation of H12 is anticipated to change from a state of enhanced free motion and distal orientation (‘open’ conformation) to a stabilized and compact orientation via interactions with helices 3 and 4 (‘closed’ conformation),\textsuperscript{1,3,9} resulting in the formation of the coregulator binding surface (\textbf{Figure 2.1}, top).

\textbf{Figure 2.1} Schematic overview of the molecular mechanism of ER LBD coactivator (LXXLL) interaction induced by an agonistic ligand (E2), followed by target gene transcription activation (top) and an alternative path of ligand-independent ER-coactivator interaction triggered by tyrosine phosphorylation (P) (bottom). Both pathways are starting from the ligand-free ‘open’ conformation and induce each an active state, represented here by a ‘closed’ conformation of H12. Such agonistic conformation is associated with instating a coactivator binding surface (light blue) on the LBD.

Coregulator proteins may have enhancing or repressing effects on target gene transcription for example by being engaged in the recruitment of enzymes which catalyze histone deacetylation or by recruiting the transcription machinery.\textsuperscript{10} An extensively studied
example of the NR-coregulator interaction is the steroid receptor coactivator 1 (SRC1) interacting with the LBD via a conserved helical LXXLL NR box motif.\textsuperscript{11–14} The common model for this protein-protein interaction (PPI) states that an agonistic closed conformation of H12 induces the formation of a ‘charge clamp’ and a hydrophobic groove,\textsuperscript{8,15,16} to which the coactivator binds via the electrostatic dipole of its helical interaction motif as well as by the accommodation of the hydrophobic leucine-residues, respectively (Figure 2.1).

The orientation of H12 is ligand-dependent. The antagonist 4-hydroxy-tamoxifen (OHT) for example induces a repressive closed conformation, which has been revealed by crystallographic studies.\textsuperscript{3} In the context of the crystal packing, OHT forces H12 to position in the ER cofactor binding groove. The complex network of PPIs, involved in ER regulation is additionally regulated by ER PTMs, such as phosphorylation and acetylation.\textsuperscript{17–24} Of specific interest in this respect are tyrosine 537 (Y537) in ER\textsubscript{α} and the analogous tyrosine 488 (Y488) in ER\textsubscript{β} which undergo context-specific\textsuperscript{18} ligand-independent phosphorylation by Src-kinase\textsuperscript{25–27} and are located at the loop-helix transition - or ‘capping motif’\textsuperscript{28} - immediately N-terminal of H12.\textsuperscript{1,29} Phosphorylation of Y537 and Y488 may therefore exert an important contribution to the dynamics of this C-terminal region and thereby have regulatory influence on ER activity\textsuperscript{28} (Figure 2.1, bottom). Different mutations of these tyrosine residues or of the amino acid residues in the immediate vicinity have already been shown to have dramatic effects on H12 dynamics and ER function\textsuperscript{28,30–39}. Furthermore, the phosphorylation status of this tyrosine has been implied to be of functional importance for the ER with respect to ligand binding affinity and kinetics, DNA binding, transactivation and PPIs.\textsuperscript{25–27,38,40–44} Mutagenesis studies of the tyrosine to aspartic and glutamic acid have shown a significantly increased activity of the ER mutants in the absence of ligand and even render the receptor consecutively active.\textsuperscript{45–47,38,48–50} Intriguingly, a number of clinical studies have recently revealed ER\textsubscript{α} Y537 mutations to be activating in hormone-resistant metastatic breast cancer.\textsuperscript{49,51–55} These important insights have mainly been obtained by cell-based and clinical studies. The exact molecular mechanism which underlies the final cellular effects has remained unresolved. Several studies have indicated that H12 dynamics and structural integrity underlie a fine-balanced regulation between active and inactive states.\textsuperscript{1,8,38,56–59} Molecular elucidation of the effect of tyrosine phosphorylation on H12 dynamics regulation and coactivator binding is therefore of great interest, regarding the clinically reported resistance inducing mutations,\textsuperscript{49,51–55} overexpression of tyrosine kinases responsible for ER phosphorylation in cancer cells,\textsuperscript{60–63} and new ligand development. A major limitation for
performing a molecular analysis into the effects of tyrosine phosphorylation on ER structure and function has been the inaccessibility of well-defined phosphorylated ER constructs for biochemical and structural studies.

This chapter covers the generation of well-defined ERα and ERβ constructs featuring either a phosphorylated or non-phosphorylated Y537/Y488 (hereafter summed up as (p)ER) in combination with site-specifically introduced fluorescent probes or 15N-labeled amino acids and their usage in elucidating the molecular role of the tyrosine phosphorylation. The introduction of the tyrosine modification as well as the read-out labels for biophysical assays are aimed for by using an efficient protein semi-synthesis approach (Figure 2.2). The introduction of the different biophysical labels is intended for the analysis of the impact of tyrosine phosphorylation on H12 dynamics and local structural arrangements by fluorescence polarization (FP), circular dichroism (CD) and NMR spectroscopy.

**Figure 2.2** Left: Image of the crystal structure of the tyrosine 488 - phosphorylated semi-synthetic ERβ LBD in complex with E2 (black, sticks) and the coactivator peptide (2) (blue, cartoon). The structure is adapted from PDB file 3OLL.29 The cartoon representation of the pY-ERβ LBD crystal structure, indicates that the ligation site lies as the junction of H11 and the loop between H11 and H12 (left). The same color code is used for the cartoon representation of the semi-synthetically produced pY-ERβ LBD (left)29 as well as the overview the NCL reaction mechanism (right). Right: Reaction mechanism of the native chemical ligation reaction between the protein thioester moiety (ER LBD, gray) and the H12 (phospho-) peptide (green), which is either unlabeled or carries labels for biophysical read-out, e.g. a fluorescent probe, linked to the side chain of C-terminal lysine or 15N-enriched amino acids.
2.2 Results and discussion

In order to study the effect of tyrosine phosphorylation on ER LBD and H12 structure, dynamics and cofactor interaction behavior, an expressed protein ligation strategy was devised,65–67 aimed to obtain a library of ER constructs featuring a phosphorylated tyrosine and molecular probes (Figure 2.3) for biophysical studies. For this purpose, the differently modified peptides were synthesized by solid phase peptide synthesis.

![Figure 2.3](image)

Figure 2.3 Library of designed semi-synthetic ER ligand binding domain constructs with (non-) phosphorylated ERα and ERβ LBDs without additional read-out labels (A), with a C-terminal fluorophore label (B, ‘F’), and (non-) phosphorylated ERβ LBD with 15N-labeled amino acid residues (C).

The C-terminal peptide moieties, intended for the introduction of the phosphotyrosines by NCL (Figure 2.2 and Figure 2.3), comprise the loop between helix 11 and 12 (L11) and H12 itself. The different peptides were synthesized by standard Fmoc-solid phase peptide synthesis.68–70 ERα and ERβ C-terminal peptides were synthesized with and without tyrosine phosphorylation, as well as featuring a site-directed insertion of a C-terminal fluorophore (fluorescein, FAM) or 15N-labeled amino acids to probe H12 dynamics by fluorescence polarization and NMR 15N-relaxation studies (Table 2.1).30
Table 2.1 (Un-) modified peptides of the C-terminal ERα and ERβ LBD moiety with or without the phospho-tyrosine (pY) of interest and including different labels for biophysical read-out methods, with site-directed introduction of fluorescein (FAM) and 15N-labeled amino acids (bold and underlined).

<table>
<thead>
<tr>
<th>Peptide name</th>
<th>Peptide Sequence</th>
<th>Calculated mass (Da)</th>
<th>Observed mass (Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td>αH12-Y</td>
<td>H-CKNVVPLYDLLLEMLDAHRLHAPT-CONH2</td>
<td>2761</td>
<td>2762</td>
</tr>
<tr>
<td>αH12-pY</td>
<td>H-CKNVVPLpYDLLLEMLDAHRLHAPT-CONH2</td>
<td>2841</td>
<td>2842</td>
</tr>
<tr>
<td>αH12-Y-FAM</td>
<td>H-CKNVVPLYDLLLEMLDAHRLK(FAM)-CONH2</td>
<td>2841</td>
<td>2841</td>
</tr>
<tr>
<td>αH12-pY-FAM</td>
<td>H-CKNVVPLpYDLLLEMLDAHRLK(FAM)-CONH2</td>
<td>2921</td>
<td>2921</td>
</tr>
<tr>
<td>βH12-Y</td>
<td>H-CKNVVPVYDLLLEMLNAHVLRG-COOH</td>
<td>2496</td>
<td>2496</td>
</tr>
<tr>
<td>βH12-pY</td>
<td>H-CKNVVPVPYDLLLEMLNAHVLRG-COOH</td>
<td>2576</td>
<td>2576</td>
</tr>
<tr>
<td>βH12-Y-FAM</td>
<td>H-CKNVVPVYDLLLEMLNAHVLRK(FAM)-CONH2</td>
<td>2769</td>
<td>2771</td>
</tr>
<tr>
<td>βH12-pY-FAM</td>
<td>H-CKNVVPVPYDLLLEMLNAHVLRK(FAM)-CONH2</td>
<td>2849</td>
<td>2848</td>
</tr>
<tr>
<td>βH12-Y-15N</td>
<td>H-CKNVVPVpYDLLLEMLNAHVLRG-COOH</td>
<td>2501</td>
<td>2500</td>
</tr>
<tr>
<td>βH12-pY-15N</td>
<td>H-CKNVVPVPYDLLLEMLNAHVLRG-COOH</td>
<td>2581</td>
<td>2580</td>
</tr>
</tbody>
</table>

2.2.1 Protein semi-synthesis of ERα and ERβ ligand binding domains

ER protein thioester constructs, constituting helices 1 until 11 of the ERα and β (K302-K529 and L260-K480, respectively) were recombinantly expressed with a C-terminal thioester, based on previously reported protocols.29,64 One-step ERα and ERβ LBD thioester affinity purification resulted in reasonably pure protein (Figure 2.4, lane 2) at a yield of about 15 mg per liter expression medium. Protein sequence integrity and thioester-stability were confirmed by QTOF LCMS analysis (Figure 2.5 A-C). Subsequent native chemical ligation between the protein thioester moiety and the synthetic peptide was performed as reported in literature with some changes to optimize the yield of the reaction.29,64 The ligation conditions were optimized for a shorter reaction time of one-time overnight incubation in view of protein stability and the excess of peptide over protein was reduced to a threefold molar excess while maintaining high yields of ligation completion (Figure 2.4).
Figure 2.4 Representative quantitative SDS-PAGE analysis (4-20 % polyacrylamide gradient gel) of the ER expressed protein thioester after one-step affinity purification of 28.0 kD (lane 2) and the native chemical ligation products ERαLBD-H12-Y-FAM of 30.7 kD (lane 3) and ERαLBD-H12-pY-FAM of 30.8 kD (lane 4) before SEC purification. Lane 5 contains a 1:1 mixture of protein thioester (1) and ligation product ERαLBD-H12-Y-FAM (3) and lane 6 contains a 1:1 mixture of protein thioester (1) and ligation product ERαLBD-H12-pY-FAM (4). Final protein samples, as used for the biophysical protein characterization are applied in lane 7 and 8 (ERαLBD-H12-Y-FAM and ERαLBD-H12-pY-FAM, respectively). A protein molecular weight marker is included in lane 1.

SDS-page analysis of the NCL reaction reveals a protein band shift from the lower molecular weight protein thioester (Figure 2.4, lane 2) to the 10 % larger ligation products (Figure 2.4, lanes 3 and 4). Gel image analysis of the ligation products before SEC purification reveals a thioester conversion of above 90 %, which represents a typical yield for this reaction for all semi-synthetically produced proteins described here (Table 2.2). The mixture of unligated and ligated protein sample confirmed the significant band-shift upon ligation (Figure 2.4, lane 5 and 6). The highly similar intensities of the protein bands corresponding to the unligated and ligated protein furthermore serve as a measure of the quantitative nature of the SDS-samples (Figure 2.4, lanes 5 and 6). Protein purity of the ligation products after SEC purification was also confirmed by this SDS-PAGE analysis (Figure 2.4, lanes 7 and 8). The latter was also confirmed by QTOF LCMS analysis of the final protein samples, as exemplified in Figure 2.5 D-F.
Figure 2.5 Representative QTOF-LCMS analysis of the semi-synthetically produced protein constructs, exemplified by ERα expressed protein thioester after one-step affinity purification (expected mass 28006 Da) with the LC-trace (A) and the m/z-spectrum (B) of the protein-thioester peak (marked by * in A and range underlined in A), including the deconvolution of the thioester peak m/z-spectrum, yielding the thioester mass in C. Right column: LC-trace (D) of SEC-purified NCL product ERαLBD-H12-pY-FAM (expected mass 30787 Da) with the the m/z-spectrum (E) of the ligation product peak (marked by * in D and range is underlined in D), including the deconvolution of the ligation product peak m/z-spectrum, yielding the product mass in F.
Taken together, the different ligation reactions successfully yielded the envisioned library of 12 different ER constructs (Figure 2.3 and Table 2.2).

**Table 2.2** Products of the protein semi-synthesis with and without phospho-tyrosine (pY), fluorescein (FAM) and 15N-labeled amino acids.

<table>
<thead>
<tr>
<th>Protein Name</th>
<th>Expected mass (Da)</th>
<th>Observed mass (Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ERα Ligation products</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ERαLBD-H12-Y</td>
<td>30645</td>
<td>Confirmed by SDS-PAGE</td>
</tr>
<tr>
<td>ERαLBD-H12-pY</td>
<td>30723</td>
<td>Confirmed by SDS-PAGE</td>
</tr>
<tr>
<td>ERαLBD-H12-Y-FAM</td>
<td>30707</td>
<td>30703</td>
</tr>
<tr>
<td>ERαLBD-H12-pY-FAM</td>
<td>30787</td>
<td>30785</td>
</tr>
<tr>
<td>ERαLBD-H12_{pren}-Y-FAM</td>
<td>30635</td>
<td>30635</td>
</tr>
<tr>
<td>ERαLBD-H12_{pren}-pY-FAM</td>
<td>30715</td>
<td>30715</td>
</tr>
<tr>
<td><strong>ERβ Ligation products</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ERβLBD-H12-Y</td>
<td>29266</td>
<td>29261</td>
</tr>
<tr>
<td>ERβLBD-H12-pY</td>
<td>29346</td>
<td>29493 (MESNA-adduct)</td>
</tr>
<tr>
<td>ERβLBD-H12-Y-FAM</td>
<td>29538</td>
<td>29538</td>
</tr>
<tr>
<td>ERβLBD-H12-pY-FAM</td>
<td>29618</td>
<td>29616</td>
</tr>
<tr>
<td>ERβLBD-H12-Y-15N</td>
<td>29270</td>
<td>29268</td>
</tr>
<tr>
<td>ERβLBD-H12-pY-15N</td>
<td>29350</td>
<td>29347</td>
</tr>
</tbody>
</table>

2.2.2 Quantification of the tyrosine phosphorylation effects on H12 dynamics and cofactor interaction

Fluorescence Polarization is a powerful method to quantify molecular interactions and elucidate structural rearrangements. It has seen application in NRs on PPARγ-LBD as well as full length ERα. In this FP assay, binding events of ligand or coactivator can influence H12 positioning, orientation, and dynamics, resulting in changes in the fluorophore polarization. Low polarization indicates a more flexible orientation of the H12 region and an increased polarization is correlated to a H12 region rigidification, in line with a ‘closed’ conformation of the LBD (Figure 2.1). Coactivator peptide titrations to the C-terminal fluorescently labeled ERα and ERβ constructs (ERα/βLBD-H12-(p)Y-FAM) were performed in order to quantify the impact of the tyrosine phosphorylation on H12 flexibility and on coactivator binding affinity. Five different coactivator peptides, based on native NR coregulators, were used to probe for potential sequence-dependent effects (Table 2.3). The coactivator peptide sequences (Table 2.3 (1)-(5)) are based on coactivators from the
Ligand-independent activation of estrogen receptor by a single PTM

extensively studied steroid receptor coactivator (SRC) family (NCOA1 and NCOA2),\textsuperscript{72} the PPARγ-coactivator 1 (PGC-1),\textsuperscript{73,74} DAX-1 (NR0B1),\textsuperscript{75,76} and RIP140 (NRIP1)\textsuperscript{77–79} and have all been shown to be involved in cellular ER pathways.

Table 2.3 Sequences and origin of the coactivator peptides, used for FP titrations.

<table>
<thead>
<tr>
<th>Peptide number</th>
<th>Peptide sequence\textsuperscript{i}</th>
<th>Protein\textsuperscript{ii}</th>
<th>Calculated mass (Da)</th>
<th>Observed mass (Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>PRQGSLYSMLTSAKQT</td>
<td>NR0B1</td>
<td>1880.2</td>
<td>1879.5</td>
</tr>
<tr>
<td>(2)</td>
<td>LTERHKILHRLLQEGSPSD</td>
<td>NCOA1</td>
<td>2228.5</td>
<td>2229.2</td>
</tr>
<tr>
<td>(3)</td>
<td>SKGQTKLLQLLTKSDQ</td>
<td>NCOA2</td>
<td>1888.0</td>
<td>1888.0</td>
</tr>
<tr>
<td>(4)</td>
<td>AEEFSLLKKLLAPANT</td>
<td>PGC-1</td>
<td>1806.0</td>
<td>1806.0</td>
</tr>
<tr>
<td>(5)</td>
<td>AANNSLLLHLLKSQTIP</td>
<td>NRIP1</td>
<td>1832.0</td>
<td>1832.0</td>
</tr>
</tbody>
</table>

\textsuperscript{i} LXXLL motif is displayed in bold; \textsuperscript{ii} native protein origin of the peptide sequence.

Experiments were performed both in the absence and presence of E2, as well as in phosphorylated and non-phosphorylated state to dissect the individual contributions of the ligand binding and tyrosine phosphorylation event (Figure 2.6 and Figure 2.7). The serial coactivator peptide titrations revealed that the presence of E2 generally lowers the \( K_D \) values of the ER variants for the interaction with the tested coactivator peptides (Figure 2.6 and Figure 2.7, black filled symbols). This observation is in line with the general consensus that agonist binding stabilizes the agonistic conformation of the protein\textsuperscript{80} and enhances the affinity for LXXLL motifs (Figure 2.1).\textsuperscript{12} The effect of tyrosine phosphorylation on cofactor binding affinities is less predictable than the described ligand-effects. It has been shown that the phosphorylation of interest does affect \textit{in vitro} evolution of LXXLL binding motifs,\textsuperscript{64} but its potential influence on naturally occurring PPIs has not been quantified before.

When comparing the peptide binding affinities of ligand-saturated ER\( \alpha \) in its phosphorylated and non-phosphorylated state (Figure 2.6 and Table 2.4 ER\( \alpha \) +E2 vs. pER\( \alpha \) + E2), no significant differences in peptide binding affinities were observed. Small modulatory effects of tyrosine phosphorylation on peptide binding affinities were detected in the comparison of these two proteins in the absence of ligand (Figure 2.6 and Table 2.4 ER\( \alpha \) - E2 vs. pER\( \alpha \) - E2). Phosphorylation slightly increased the affinity of ER\( \alpha \) - E2 in the case of peptide (1) and decreased the affinities for peptide (2) and (4). Remarkably, in contrast to ER\( \alpha \), phosphorylation of ER\( \beta \) in the absence of E2 strongly increased the affinity for all tested coactivator peptides significantly (Figure 2.7 and Table 2.4 ER\( \beta \) - E2 vs. pER\( \beta \) - E2).
Figure 2.6 Fluorescence polarization data obtained from titrations of ERαLBD-H12-(p)Y-FAM, here abbreviated as (p)ERα, with coactivator peptides (1), (2) and (4) (A, B and C, respectively). Titrations are performed with phosphorylated and non-phosphorylated LBDs (circles and squares, respectively) as well as in the presence (filled symbols) and absence (empty symbols) of agonistic ligand 17β-estradiol (E2, black) and antagonist 4-hydroxy-tamoxifen (OHT, gray). Concentration of FAM-ERα was kept constant at 100 nM. E2 as well as OHT were added at a constant concentration of 10 µM resulting in a final concentration of 0.05 % DMSO, which was also included in the ligand-free samples.

The observed impact of the ERβ phosphorylation reaches up to the same order of magnitude as the effect of E2 on peptide binding affinities. The observation that phosphorylation already caused high binding affinities in the absence of E2, brings forth the expectation that the added effect of E2 will be diminished. This is in line with the small enhancement of pERβ binding affinities upon addition of E2 (Figure 2.7 and Table 2.4 pERβ - E2 vs. pERβ + E2).
Figure 2.7 Fluorescence polarization binding curves obtained from titrations of FAM-(p)ERβ LBD with coactivator peptides (1)-(5) (A until D, respectively). Titrations are performed with phosphorylated and non-phosphorylated LBDs (circles and squares, respectively) as well as in the presence (filled symbols) and absence (empty symbols) of agonistic ligand 17β-estradiol (E2). Concentration of FAM-ERβ was kept constant at 25 nM. E2 was added at a constant concentration of 10 μM resulting in a final concentration of 0.05 % DMSO, which was also included in the ligand-free samples.
Table 2.4 $K_D$ values (µM) for the binding of coregulator peptides to fluorescently labeled phosphorylated or non-phosphorylated ERα and ERβ constructs in the absence or presence of E2.

<table>
<thead>
<tr>
<th>Protein (Peptide)</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERα - E2</td>
<td>6.1 ± 0.9</td>
<td>2.0 ± 0.5</td>
<td>n.d.</td>
<td>61.7 ± 0.4</td>
<td>n.d.</td>
</tr>
<tr>
<td>ERα + E2</td>
<td>0.4 ± 0.04</td>
<td>0.2 ± 0.04</td>
<td>n.d.</td>
<td>3.6 ± 0.7</td>
<td>n.d.</td>
</tr>
<tr>
<td>pERα - E2</td>
<td>4.4 ± 0.05</td>
<td>3.7 ± 0.6</td>
<td>n.d.</td>
<td>101 ± 9</td>
<td>n.d.</td>
</tr>
<tr>
<td>pERα + E2</td>
<td>0.3 ± 0.04</td>
<td>0.6 ± 0.1</td>
<td>n.d.</td>
<td>3.2 ± 0.4</td>
<td>n.d.</td>
</tr>
<tr>
<td>ERβ - E2</td>
<td>1.8 ± 0.1</td>
<td>14 ± 1</td>
<td>13 ± 1</td>
<td>26 ± 2</td>
<td>6.9 ± 0.4</td>
</tr>
<tr>
<td>ERβ + E2</td>
<td>0.030 ± 0.001</td>
<td>0.18 ± 0.02</td>
<td>0.12 ± 0.02</td>
<td>0.24 ± 0.05</td>
<td>0.05 ± 0.01</td>
</tr>
<tr>
<td>pERβ - E2</td>
<td>0.08 ± 0.01</td>
<td>0.24 ± 0.02</td>
<td>0.32 ± 0.02</td>
<td>0.76 ± 0.07</td>
<td>0.16 ± 0.02</td>
</tr>
<tr>
<td>pERβ + E2</td>
<td>0.025 ± 0.004</td>
<td>0.13 ± 0.03</td>
<td>0.07 ± 0.01</td>
<td>0.09 ± 0.01</td>
<td>0.05 ± 0.01</td>
</tr>
</tbody>
</table>

n.d. = not determined.

Altogether, tyrosine phosphorylation exerts an ER subtype specific effect on the tested interactions. In the case of ERβ, tyrosine phosphorylation alone is sufficient to introduce agonist-like affinities towards the coactivator-peptides. Whereas in the case of ERα, phosphorylation of the tyrosine does not induce high ligand-independent coactivator affinities and also does not appear to affect the binding modes uniformly in the presence of agonist. The observed effect of phosphorylation on ERβ coactivator binding behavior, suggests that phosphorylation of Y488 induces conformational or dynamic alterations that result in a shift of the conformer-equilibrium towards an agonist-like state in the absence of ligand. Insights on the underlying structural and dynamic alterations can be obtained directly from the observed fluorescence polarization data by comparing the polarization values at low coactivator peptide concentrations (Figure 2.6 and Figure 2.7).

Based on the FP theory,$^{71}$ high starting levels of polarization (200 ± 10 mP) presumably confer to a more rigid closed conformation of H12 and low starting levels of polarization (160 ± 10 mP) to a more flexible open H12 conformation (Figure 2.1). This assumption was supported by experiments with OHT, which forces H12 in an open conformation, resulting in polarization values slightly lower than 160 mP (Figure 2.6). Polarization levels of the different ERβ variants at low cofactor concentrations (Figure 2.7) revealed low polarization levels only for the ligand-free (p)ERβ (Figure 2.7, empty symbols). High starting levels of polarization were observed in the presence of E2 for both phosphorylated and non-phosphorylated ERβ (Figure 2.7, filled symbols). This observation indicates, that ERβ-H12 adopts a closed conformation in the presence of E2. Interestingly, the pERβ H12 resides in an open conformation in the absence of ligand (Figure 2.7, empty...
circles). However, tyrosine phosphorylation apparently allows the transition towards the closed conformation at lower coactivator peptide concentrations (Figure 2.7, empty circles vs. empty squares). In contrast to ERβ H12 behavior, a different pattern was observed for the polarization levels of ERα H12 at low coactivator peptide concentrations. The starting polarization values of ligand-bound as well as ligand-free ERα and ligand-free pERα resided around 160 mP (Figure 2.6, filled squares and empty circles, respectively). Only the ligand-bound pERα exhibited high starting polarization levels (Figure 2.6, filled circles). Consequently, ERα-H12 appears to be in a highly dynamic open state also in the presence of agonist. Only upon phosphorylation, ligand-saturated ERα-H12 adopts a closed conformation. These observations add a level of complexity to the widely excepted activation model of ER, since it appears that saturation of the ligand binding site with E2 does not imperatively induce a closed conformation to ERα. Taken together, tyrosine phosphorylation appears to exert a stabilizing effect on both, ERα and ERβ. The effect on ERβ is translated to higher binding affinities towards the tested coactivator peptides, whereas the effect of ERα-phosphorylation is only visible in the ligand-bound state and therefore presumably does not add to the actual affinities towards the tested coactivator interactions. The modulatory effect of the phospho-group is also reflected in the full ligand and coactivator bound state as highlighted by the different end-levels of the polarization values. Interestingly, in almost all of the tested interactions these end-levels are the same for each phosphorylated or non-phosphorylated construct in the absence or presence of E2, but differ depending on their phosphorylation status. Also the cofactor-saturated agonistic solution conformation is thus fine-tuned by the phosphorylation status. This observation is in line with the differential cofactor-preferences which previously evolved for the phosphorylated and non-phosphorylated ER in a cofactor in vitro evolution approach.

In order to further elucidate the subtype-specific contributions of the semisynthetically introduced C-terminal H12-loop region (Figure 2.2), ERα LBD core (H1-H11) was ligated to the ERβ C-terminal moieties βH12-Y-FAM and βH12-pY-FAM, individually (Table 2.1). Interestingly, binding affinities of the resulting protein constructs ERαLBD-H12β-Y-FAM and ERαLBD-H12β-pY-FAM (Table 2.2) to peptides (2) and (4) were strongly decreased (data not shown), indicating an impairment of H12-function within the altered protein context. The binding affinity towards peptide (1), however, remained within the accessible concentration range of the FP assay (Figure 2.8).
The comparison between the FP assay of the mixed constructs (p)ERα-H12\textsubscript{beta} with the earlier reported results of the homogenous protein constructs (Figure 2.6 until Figure 2.8 and Table 2.4) revealed that the (p)ERβ protein constructs bind peptide (1) with significantly higher affinity than the ERα protein constructs under the tested conditions (Figure 2.8 and Table 2.5). Furthermore, the comparison of each binding curve for the mixed ligation constructs ((p)ERα-H12\textsubscript{beta}) with its respective ERα equivalent ((p)ERα), reveals smaller differences in the binding behavior towards peptide (1) than in comparison to the respective ERβ construct. Especially the binding affinity, determined for the phosphorylated and ligand saturated mixed ERα protein construct is very similar to the binding affinity determined for

**Figure 2.8** Fluorescence polarization assay of the protein constructs ERαLBD-H12\textsubscript{beta}-pY-FAM, here abbreviated as (p)ERα-H12\textsubscript{beta}, (black triangles) in comparison to the binding curves of phosphorylated (circles) and non-phosphorylated (squares) ERα (black) and (p)ERβ (gray). All titrations are performed with peptide (1). Binding curves were obtained in the absence of ligand E2 (empty symbols A and C) and in the presence of E2 (B and D) as well as in the non-phosphorylated state (A and B) and in the phosphorylated state (C and D). Concentration of (p)ERα-H12\textsubscript{beta} constructs was kept constant at 100 nM. E2 was added at a constant concentration of 10 µM resulting in a final concentration of 0.05 % DMSO, which was also included in the ligand-free samples. The titration curves for (p)ER and (p)ER were included from data-sets, described in Figure 2.6 and Figure 2.7.
the homogeneous pERα construct (Figure 2.8 D, Table 2.5 and for comparison Table 2.4). The differences between the non-phosphorylated ERα-constructs (Figure 2.8 A and B) turned out to be larger than the differences between the phosphorylated ERα-constructs (Figure 2.8 C and D), indicating that in this protein-domain context, tyrosine phosphorylation makes the ERβ-C-terminal sequence adapt a more ERα-like state and contribute in such way to the coactivator interaction. The differences in starting and saturation levels in fluorescence polarization of the individual samples upon introduction of the H12beta-sequence into the ERα LBD does not provide additional information on how the changes in binding affinities are mediated via e.g. altered H12-dynamics and orientation (Figure 2.8).

Table 2.5 \(K_D\) values (\(\mu\)M) for the interaction between (non-) phosphorylated (p)ERα-H12beta protein constructs and peptide (1) in the presence and absence of E2.

<table>
<thead>
<tr>
<th>Peptide (1)</th>
<th>ERα-H12beta - E2</th>
<th>ERα-H12beta + E2</th>
<th>pERα-H12beta - E2</th>
<th>pERα-H12beta + E2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peptide (1)</td>
<td>52.5 ± 6</td>
<td>1.3 ± 0.2</td>
<td>18 ± 1</td>
<td>0.6 ± 0.1</td>
</tr>
</tbody>
</table>

The quantitative analysis of the binding curves determined for the mixed (p)ERα-H12beta constructs reveals that not only the obvious differences with respect to the (p)ERβ samples are noteworthy, but that also the differences in binding affinities with respect to the (p)ERα constructs are significant (Table 2.4 and Table 2.5). In fact, the introduction of the ERβ-derived C-terminal sequence into the ERα LBD results in 2-3 fold increase in the \(K_D\) towards peptide (1) in the presence of E2 and in a 4-9 fold increase in the absence of E2 (Table 2.4 and Table 2.5).

Taken together, these results confirm the mediation of the cofactor binding event by a combination of the LBD core and the C-terminal region and its fine regulation by sequence intrinsic protein characteristics. The pronounced differences between the wild type and the mixed constructs furthermore emphasize the remarkably high regulatory potential of the C-terminal region, since the observed altered behavior is induced by only 3 differences in the amino acid sequence (Table 2.1). Furthermore the phosphorylation in combination with E2 can recover the ERα wild type binding behavior to the tested peptide, pointing out the strength of the tyrosine phosphorylation effect in ER LBD activity regulation.
2.2.3 On-chip cofactor-interaction screening

The influence of the receptor phosphorylation on its binding behavior to a large set of cofactor peptides, derived from natural cofactor sequences, has been investigated in an on-chip binding array (Figure 2.9). In this array, 53 naturally derived cofactor peptide sequences are immobilized in individual spots, allowing the parallel screening of all different sequences under identical experimental circumstances. The on-chip array was performed in the absence and presence of E2 for each protein subtype and in the phosphorylated as well as non-phosphorylated state. An overview of the resulting relative differences in binding levels, induced by either the ligand or the phosphorylation is given in Figure 2.9.

Figure 2.9 Relative changes in ERα (left column) and ERβ (right column) binding levels towards an on-chip array of 53 coregulator sequences, upon addition of E2 (A-D) and upon phosphorylation (E-H). Changes in binding levels as response to E2 by ERα (A) and ERβ (B) and to the phosphorylated pERα (C) and pERβ (D) as well as changes induced by phosphorylation of ERα and ERβ in the absence of E2 (E and F, respectively) and in the presence of E2 (G and H, respectively).
The summary of relative changes in ER-binding levels to the cofactor arrays revealed a strong increase in ER binding levels to the majority of the peptides upon addition of E2 to both ERα as well as ERβ (Figure 2.9 A and B, respectively). Furthermore the response of pERα to E2 (Figure 2.9 C) appeared stronger for individual coactivator sequences than for non-phosphorylated ERα. This is in line with the induced closed conformation of ERα by phosphorylation in the presence of E2 (Figure 2.6 A, filled circles vs. empty circles). In contrast, the changes of pERβ binding levels upon addition of E2 are negligible (Figure 2.9 D), which is in line with the small differences in binding affinities between pERβ + E2 vs. ERβ + E2, observed in the FP assay (Table 2.4). The overall changes of binding levels upon phosphorylation of ERα and ERβ (Figure 2.9 E - H) also compare well to the effects of phosphorylation, observed in the FP assay. In the absence of E2, phosphorylation of ERα has no significant effect on the binding levels (Figure 2.9 E). In contrast, the overall increase in binding levels, induced by phosphorylation of ERβ in the absence of E2 is high (Figure 2.9 F). This reflects the observation of higher binding affinities of ERβ - E2 upon phosphorylation in the FP assay (Table 2.4) and supports the hypothesis of a ligand-independent subtype specific activation mechanism for ERβ-targeted gene expression. The effect of phosphorylation on ERα interaction with the immobilized peptides in the presence of E2, also revealed a remarkable binding behavior (Figure 2.9 G). Next to a small overall-increase of binding levels, ERα-phosphorylation in the presence of E2 induced a specific increase of binding to a set of corepressor sequences (Figure 2.9 G). Phosphorylation of ERβ in the presence of E2, however, does not further increase the binding levels (Figure 2.9 H). Taken together, the on-chip binding array revealed the overall effects of ERα and ERβ phosphorylation on a large group of native cofactor peptides. The results complement the FP results and hint at an effective cellular ER subtype specific regulation mechanism by tyrosine phosphorylation.

2.2.4 Circular dichroism of L11-H12 peptides

CD spectroscopy is a technique that can yield valuable data on the structural characteristics of proteins and polypeptides. Here, CD spectroscopy it was employed to monitor the effect of phosphorylation on the secondary structure of the isolated C-terminal protein sequences (Table 2.1), for both ERα and ERβ (H12α and H12β, respectively). Based on crystal structures of ERα and ERβ, this C-terminal region consists of a loop and a short H12 in the agonist-bound state. The CD spectra, recorded in sodium phosphate buffered
aqueous solutions of the non-phosphorylated peptides (**Figure 2.10**, dashed lines), feature a signature typical for α-helical peptides with θ_{222}/θ_{208} ratios of 0.6 and 0.9 (H12α and H12β, respectively). This indicates that this C-terminal segments harbor a considerable degree of intrinsic helicity. Tyrosine phosphorylation leads to a CD signature with a significantly lowered α-helical character (**Figure 2.10**, grey lines) for both peptides. The θ_{222}/θ_{208} ratios decreased to 0.5 for both phosphorylated sequences (pH12α and pH12β), suggesting a disruption, at least to some extent, of the initial length of H12, by the phosphorylation.

**Figure 2.10** Far-UV CD spectra of the isolated C-terminal peptides of ERα (A) and ERβ (B) in non-phosphorylated (dotted line) - H12α and H12β, respectively - and phosphorylated state (gray solid line) - pH12α and pH12β, respectively.

Phosphorylation of amino acid side chains within α-helices can have stabilizing or destabilizing effects on the secondary structure, depending on the location within the helix as well as the sequence context. The tyrosine of interest is located at the N-terminus of the H12 core motif of ERα and ERβ and is part of the H12 capping motif. Additionally, this tyrosine is part of the intramolecular H12 corepressor motif P(L/V)YDLLLEML (amino acid sequence of the ER(α/β) core motif with the corepressor motif in bold), and its molecular status will thus impact the role and conformation of H12. Other steroid NRs feature a conserved proline at the ER tyrosine position (**Figure 2.11**, * symbol). This proline contributes to the orientation of H12 and acts as H12 capping residue. In ER, the tyrosine thus takes over the role of the proline cap and the effectiveness thereof is potentially regulated via its phosphorylation state. ERα and ERβ actually feature a proline residue 2 positions more N-terminal of the tyrosine (**Figure 2.11**). The different location of this proline with respect to the H12 core motif of the other steroid NRs implies differences in the (non-) phosphorylated H12 length and consequently differences in helicity of the intramolecular cofactor binding motifs. The general importance thereof has been associated with different cofactor interaction tendencies.
Figure 2.11 Sequence alignment of amino acids of the loop region preceding H12 for several human steroid NRs and hRXRa with the H12 capping residues indicated by an asterisk (*) and the different positions of conserved proline residues highlighted by black boxes.

The enhanced binding of corepressor sequences in the case of ligand-bound phosphorylated ERα (Figure 2.9 G) supports the hypothesis that the competition of the corepressor binding motif, located in H12, with the corepressor sequences is reduced by disruption of the helicity of the internal corepressor motif in H12. Crystal structures of ligand-free ER are not available. However, in the ligand-free structure of the NR RXR, H11 extends into H12 without being interceded by a flexible loop. Remarkably, RXR and ER both feature the aforementioned common proline at the same upstream position (Figure 2.11). This implies that in the unliganded state the H12 of ER extends until this proline and that tyrosine phosphorylation shortens the length of H12. This shorter H12 is typically observed in the agonistic state seen in the ligand bound crystal structure and in line with the position and effect of the proline in the other steroid NRs. Ligand binding may thus induce a partial unfolding of the H12 helical segment, to enable the shortened H12 to fold back on to the LBD. The tyrosine phosphorylation facilitates or enhances this process and thus directs the equilibrium of ER conformers to an agonistic state.

2.2.5 Solution NMR of the 15N-labeled ERβ LBD

A well-established method to gain insight on protein dynamics is solution NMR. The quality of protein NMR spectra is limited by their resolution, which is affected by the size of the protein of interest. The signal-to-noise quality of protein-NMR spectra may generally be improved by longer recording times, which requires high protein stability under experimental conditions. So far, no NMR data has been reported for the ER LBD due to these
limitations. The site-directed introduction of $^{15}\text{N}$-labeled amino acid residues by the described semi-synthesis approach allows to probe phosphorylation-induced changes in the C-terminal protein region specifically (Figure 2.3). $^{15}\text{N}$-labels were introduced throughout the C-terminal $\text{ER}_\beta$ LBD region covering the sequence, corresponding to the loop between H11 and H12 as well as the sequence, corresponding to H12, which serve as probes for dynamic changes within the loop region ($^{15}\text{N}$-K482), in the direct vicinity to the tyrosine of interest ($^{15}\text{N}$-V487) and within the H12 sequence ($^{15}\text{N}$-A497) (Table 2.1). The $\text{ER}_\beta$ subtype constructs were chosen based on the observed subtype-specific phosphorylation effects and the observation that $\text{ER}_\beta$ LBD protein constructs are more stable under the experimental conditions in comparison to $\text{ER}_\alpha$ LBD protein constructs. Eight different protein samples were prepared for NMR studies, covering the non-phosphorylated and phosphorylated $\text{ER}_\beta$ in the free state, in the presence of either E2 or peptide (2), and in the presence of both E2 and peptide (2).

**Figure 2.12** A) and B) Comparison of the $^{1}\text{H}$$-^{15}\text{N}$-HSQC spectra of $\text{ER}_\beta$ (A) and p$\text{ER}_\beta$ (B) in the free state (black), in the presence of peptide (2) (red), in the presence of estradiol (green), and in the presence of E2 and peptide (2) (blue). The upper plot shows a superposition of the four spectra and at the bottom, each spectrum is shown in the same color coding separately. Lines label peaks that may represent additional conformational states of the respective residue. Dotted lines indicate that the respective peaks could not be assigned.
The individual recorded $^1$H-$^{15}$N-HSQC spectra together with their superposition (Figure 2.12) reveal overall clear differences between the non-phosphorylated and phosphorylated ER$\beta$, depending on the sample conditions (Figure 2.12 A and B, respectively). In the $^1$H-$^{15}$N-HSQC spectra of the non-phosphorylated ER$\beta$ (Figure 2.12 A) A497 shows generally two neighboring peaks, indicating the existence of two distinct conformers in slow chemical exchange. Addition of E2 or peptide (2) resulted in significant effects on the signal intensities. Changes in chemical shift were not observed. This may be explained by a change of the dynamic properties of the C-terminal region upon addition of agonistic ligand. Binding of E2 resulted in overall lower peak intensities, suggesting H12 resonances to be in intermediate chemical exchange, which causes peak broadening. In contrast, binding of peptide alone or in combination with E2 resulted in a clear increase of V487 and A497 signal intensities, which may be explained by a stabilization of the conformation of this segment compared to the apo-ER$\beta$. This would also be consistent with the observed increase of the FP levels of the ER$\beta$ upon titration with peptide (Figure 2.6 and Figure 2.7).

The $^1$H-$^{15}$N-HSQC spectra of the phosphorylated ER$\beta$ in the four different conditions reveal a significantly different signature than the non-phosphorylated ER$\beta$ (Figure 2.12 B). In the spectrum of the apo form, a single peak is visible for A497, whereas K482 and V487 appeared to show multiple peaks. This observation indicates the presence of local structural heterogeneity in the direct vicinity of K482, which is in line with the observation that tyrosine phosphorylation disrupts the helical character of the C-terminal ER moieties (Figure 2.10). Addition of only peptide (2) resulted in even more peaks for K482 and a shoulder in the peak assigned to A497, while binding of E2 or E2 and peptide (2) simultaneously reduced the number of peaks. Importantly, the set of peaks observed for K482 in the free and peptide (2)-bound state merged into a single resonance in the complexes with either only E2 as well as in the presence of both, E2 and peptide (2). This fused peak resides at about the same position as the one observed for the non-phosphorylated protein (Figure 2.12 A). Generally, a reduction of the number of peaks to a single resonance may be explained by reduced conformational heterogeneity and thus the stronger population of a single (agonistic) conformation or an increase in the backbone dynamics such that only an average signal is visible for the fast exchanging populated conformations.

Taken together, these observations indicate, that either in the free form or in the presence of only peptide (2), phosphorylation of Y488 results in a change in dynamics and an
increased conformational heterogeneity in the loop encompassing residues K482-A487. However, upon addition of E2, this conformational heterogeneity becomes restricted, resulting in similar spectra as the ones observed for the non-phosphorylated protein \textit{i.e.} a fully agonistic conformation.
2.3 Conclusion

The successful protein semi-synthesis of a library of 12 different protein constructs, designated for the analysis by different biophysical techniques is summed up in this chapter. ER LBD semi-synthesis allowed for the introduction of the regulatory pY537 into ERα and pY488 in ERβ as well as the site-specific introduction of a fluorescent probe at the very C-terminus of the (phosphorylated) proteins and the site-directed introduction of 15N-labeled amino acids for the read out of phosphorylation effects by NMR. Furthermore the semisynthetic approach was employed to construct mixed protein constructs by introducing the C-terminal synthetic moiety of phosphorylation-regulated ERβ to the ERα LBD, allowing for an estimation of the contribution of ER LBD and the C-terminal protein region to the observed phosphorylation-effects. In addition, the intrinsic changes in structural signature of the C-terminal moiety, induced by phosphorylation, were characterized by CD spectroscopy of the synthetic C-terminal peptides. The fluorescence polarization cofactor binding assays revealed distinct effects of tyrosine phosphorylation on H12 orientation as well as changes in cofactor binding affinities in an ER subtype specific manner. An outstanding effect of tyrosine phosphorylation was observed for the strongly enhanced binding affinities of ERβ to the coactivator sequences in the absence of E2. This enhancement occurs in the order of magnitude as the E2-induced cofactor interaction, indicating an alternative phosphorylation-regulated and subtype specific ERβ activation mechanism. Intriguingly, phosphorylation of ERα appears to enhance specific corepressor interactions in the presence of E2, as indicated by an on-chip peptide binding array. Taken together, these phosphorylation effects strongly suggest an alternative ligand-independent and subtype specific regulation mechanism in which ERβ activity is directly enhanced by coactivator interaction stabilization whereas ERα activity is specifically inhibited by corepressor recruitment.

In line with the widely accepted model of the transition from an open to a stably closed conformation of H12 in the presence of E2, the ERβ LBD C-terminus was observed to reside in a state of reduced rotational freedom – a ‘closed’ state - in the presence of E2, as indicated by high starting levels of fluorescence polarization of the phosphorylated as well as the non-phosphorylated ERβ LBDs. For ERα, however, the reported data suggests an extension of the model, since such a state of reduced rotational freedom, indicating a closed an potentially more agonist conformation of H12, was only observed in the E2 saturated and phosphorylated ERα, yet not in the non-phosphorylated ERα. These observations indicate that the tyrosine phosphorylation stabilizes a closed state in ERα. Interestingly, tyrosine
phosphorylation effects are translated into ER-agonism in a subtype specific manner. While the apparent stabilization of E2-bound ERα-H12 in a closed conformation by tyrosine phosphorylation does not increase coactivator binding, tyrosine phosphorylation of ERβ in the absence of agonist induces high binding affinities to the coactivator peptides, yet no closed H12-conformation. This complexity of H12-positioning and its subtype specific translation into ER-agonism emphasizes the fine regulation of this C-terminal protein region and illustrates the potential of small modifications, such as the tyrosine phosphorylation or clinically relevant mutations within this protein region. The classical model of ER antagonism by OHT, which has been reported to induce a repressive closed conformation to H12-orientation, may also need extension away from the static picture of a repressive closed conformation. The FP data, reported here rather suggests that OHT forces H12 into an open conformation, inhibiting the formation of the coactivator binding site. The direct influence of tyrosine phosphorylation on local structural signature and dynamic behavior of the surrounding sequence was investigated by CD and NMR spectroscopy. CD spectroscopy revealed a significant reduction of the helical signature upon tyrosine phosphorylation which appears to be stronger in the case of the ERβ subtype. The observed increased structural disorder upon tyrosine phosphorylation specifically in the N-terminal region of the H12-sequence was confirmed by NMR experiments. This observation together with the role of Y537 and Y488 as H12 capping residue in the agonist-bound structural conformation of ER LBD, suggests a translation of the tyrosine phosphorylation effect via an altered structural flexibility around the phosphorylation site. On the whole, the described data suggests a model in which the ‘H11-H12’-region has a high alpha-helical character in the absence of E2 in analogy to the extended H11/H12 in ligand-free hRXRα. Tyrosine phosphorylation then exerts its regulatory character by inducing enhanced flexibility to this region, shifting the conformational state more towards the reported ‘H11-loop-H12’-secondary structure and thereby allowing the adaptation of an agonistic conformation in the presence of cofactor peptide. The observed cofactor binding behavior of the ERα LBD core ligated to the ERβ C-terminal moiety furthermore indicated that the complex mechanism of ER regulation by tyrosine phosphorylation is established by a fine balanced interplay between the LBD core and the C-terminal LBD region. Altogether, this study of the molecular effects of tyrosine phosphorylation emphasizes the potential of protein semi-synthesis approaches to deepen our understanding of fundamental protein regulatory mechanisms and the elucidation of clinically relevant protein modifications. This knowledge and the described tools hold the potential to
direct the development of new molecules, which can specifically antagonize phosphorylation up-regulated ER activity.
2.4 Experimental section

**Peptide synthesis.** All peptides were synthesized by standard Fmoc- solid phase peptide synthesis (SPPS). In general, the peptides were synthesized on Rink amide resin if not stated otherwise. An Fmoc-Tyr(PO(OBzl)OH)-OH (EMD Millipore) building block was used for the introduction of the phosphorylated tyrosine residue. Successful synthesis of all peptide sequences was confirmed by analytical LCMS. High purity of all synthesized peptides was assured by purification via a water-acetonitrile gradient in the presence of 0.1 % TFA by reversed phase HPLC with an ESI MS-detector. Peptides $\beta$H12-(p)Y and $\beta$H12-(p)Y-$^{15}$N (Table 2.1) were synthesized on Fmoc-Gly-functionalized Wang resin (EMD Millipore). The $^{15}$N-labels were introduced by coupling the minimal 98% $^{15}$N–enriched building blocks Fmoc-Lys(BOC)-OH-$^{15}$N (ISOTEC, Sigma Aldrich), L-Alanine-N-Fmoc and L-Valine-N-Fmoc (Cambridge Isotope Laboratories). The fluorescently labeled peptides $\alpha$H12-(p)Y-FAM and $\beta$H12-(p)Y-FAM (Table 2.1) were synthesized on a Tenta Gel R RAM resin (Rapp Polymere). A Dde-protected Fmoc-lysine was introduced C-terminal to the native H12 sequence. The orthogonal deprotection of the lysine side chain, allowed for coupling of 5- or 5/6-carboxy-fluorecein (FAM) to the lysine side chain NH$_2$-group in a repeated overnight reaction in the presence of 4 equivalents of FAM, 3.6 equivalents of HATU and 8 equivalents of DIPEA.

**Protein thioester expression and purification.** The sequences of ER$\alpha$ (K302-K529) and ER$\beta$ (L260-K480) were expressed from pTWIN1 vectors (New England Biolabs) in E. coli BL21(DE) cells (EMD Millipore) as fusion proteins with a C-terminal Mxe intein domain to yield the C-terminal thioester after thiol-induced intein cleavage. A C-terminal chitin binding domain was used for affinity purification on chitin-functionalized resin (New England Biolabs). An N-terminal His$_8$-tag was introduced to serve as fluorescent antibody binding epitope in the on-chip coregulator binding assay for detection. After 20 h expression at 15°C in TB medium, cells were opened by homogenization (EmulsiFlex C3, Avestin) in Buffer A (20 mM Tris at pH 7.5, 500 mM NaCl, 1 mM EDTA, 0.1 mM TCEP) supplemented with 200 µM PMSF, 1 µg/ml DNAse I (EMD Millipore), and 1 mM MgCl$_2$. Affinity purification on chitin beads at 4°C in Buffer A, was followed by on-column intein cleavage at pH 8.0-8.5 in Buffer A with 200 mM MESNA for 16 hours at 4°C. After elution in the same buffer, the cleavage step was repeated. Protein concentration was increased by centrifugation in Amicon® Ultra 10K filter devices and the protein thioester was stored in the elution buffer at -80°C until ligation to H12-peptides. Altogether, this procedure typically yields around 15 mg protein thioester per liter TB medium.

**Native chemical ligation of the different protein constructs.** Prior to the ligation reaction, the respective H12-peptide was incubated at 20°C under reducing conditions (2 mM TCEP) to ensure reduced cysteine side chains. For the ligation reactions, protein thioester was incubated with 3 equivalents of the respective H12-peptide overnight at 4 °C and pH 8.0 in Buffer A supplemented with 2 mM TCEP and 200 mM MESNA. Subsequent purification by means of size exclusion ensured the removal of unligated peptide.

**Fluorescence polarization assay.** Fluorescence polarization (mP) was measured with a Tecan Safire monochromator microplate reader. The coregulator titrations to FAM-labeled ERs were performed in TR-FRET coregulator buffer E at pH 7.5 (Invitrogen), supplemented with 1 mM DTT and either 10 µM E2 or 4-hydroxytamoxifen, dissolved in DMSO, or DMSO alone, resulting in a final DMSO concentration of 0.05%. The coregulator peptides were dissolved in coregulator buffer E likewise and serially diluted from millimolar to nanomolar concentrations. Protein concentrations were kept constant at 100 nM ((p)ER$\alpha$) and 25 nM ((p)ER$\beta$).
Fluorescence polarization was determined by recording the excitation at 470 nm and emission at 519 nm. The polarization ([mP], Y) was plotted against the coregulator peptide concentration ([log M], X). The $K_D$ values in Table 2.4 and Table 2.5 were determined by fitting the binding data with the equation $Y = B_{\text{max}} \times X / (K_D + X) + BG$. In which $B_{\text{max}}$ is the maximum binding at saturating levels of coregulator binding peptide and the polarization level at the lowest concentration of coregulator peptide is set as background (BG).

**On-chip high throughput coregulator binding array.** The on-chip array (PamGene) was performed in TR-FRET coregulator buffer E at pH 7.5 (Invitrogen) supplemented with 5 mM DTT at room temperature. 4 nM of the respective ER construct were applied to the chip in the presence or absence of 10 μM E2. The ligand was dissolved in DMSO and the ligand-free samples were supplemented with DMSO only, yielding in all cases a final DMSO concentration of 0.1%. A final concentration of Penta-His Alexa Fluore 488 antibody conjugate (Qiagen) was applied to the immobilized ER. Unspecific binding to the chip is avoided by treatment of the chip with 1% BSA- and 0.01% Tween- supplemented TBS solution.

**CD spectroscopy of (p)ERα and (p)ERβ H12 peptides.** The CD spectra were recorded in 5 mM sodium phosphate buffer at pH 7.8 and 25 °C in 1 mm path length quartz cells. At peptide concentrations of 25 μM for (p)ERα and 90 μM for (p)ERβ, 5 accumulations per sample were recorded on a Jasco J-18 CD spectrometer with a temperature controller (PTC432S).

**NMR spectroscopy.** The NMR-samples of $^{15}$N-labeled ERβ at concentrations of about 0.09-0.15 mM. $^{15}$N-labeled pERβ protein samples concentration of about 0.14 mM. NMR-samples were prepared in 20 mM Tris at pH 7.5, 300 mM NaCl, 1 - 1.4 mM TCEP and containing 5% D2O. The details of the analyzed samples are listed in Table 2.6.

**Table 2.6** Protein, peptide and ligand concentrations of the NMR samples.

<table>
<thead>
<tr>
<th>Sample</th>
<th>ERβ [mM]</th>
<th>pERβ [mM]</th>
<th>Peptide (2) [mM]</th>
<th>E2 [mM]</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERβ free</td>
<td>0.103</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ERβ + E2</td>
<td>0.086</td>
<td>0</td>
<td>0</td>
<td>0.171</td>
</tr>
<tr>
<td>ERβ + peptide (2)</td>
<td>0.096</td>
<td>0</td>
<td>0.280</td>
<td>0</td>
</tr>
<tr>
<td>ERβ + peptide (2) + E2</td>
<td>0.096</td>
<td>0</td>
<td>0.280</td>
<td>0.240</td>
</tr>
<tr>
<td>pERβ free</td>
<td>0</td>
<td>0.144</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>pERβ + E2</td>
<td>0</td>
<td>0.144</td>
<td>0</td>
<td>0.280</td>
</tr>
<tr>
<td>pERβ + peptide (2)</td>
<td>0</td>
<td>0.141</td>
<td>0.392</td>
<td>0</td>
</tr>
<tr>
<td>pERβ + peptide (2) + E2</td>
<td>0</td>
<td>0.141</td>
<td>0.392</td>
<td>0.275</td>
</tr>
</tbody>
</table>

E2 was diluted from a stock solution in DMSO. Final DMSO concentration was not exceeding 2.8 % (v/v). NMR spectra were acquired at 10 °C on Bruker Avance 500 and 750 MHz spectrometers, the 500 MHz one equipped with a cryogenic probe. The data were processed with NMRPipe and analyzed using NMRView. Assignments for the three $^{15}$N-labeled residues were based on 3D $^{15}$N-edited NOESY-HSQC spectra.
2.5 Acknowledgements

Part of the PamGene experiments as well as the cloning of the protein constructs and their use in protein semi-synthesis has been described previously in the PhD thesis of Sabine Moecklinghoff (Modulating the Nuclear Receptor-Cofactor Interaction, Characterization and Inhibition). The protocol for the peptide synthesis of the fluorescently labeled H12 was established by Christian Haase during the synthesis of the FAM-ERβ-peptides and later applied by Mark Balk to the FAM-ERα-peptides. The coactivator peptides were synthesized by Marcel Scheepstra. Matthew Burton helped with the purification of the $^{15}$N-labeled H12 peptides. Wencke Adriaens established the initial cofactor titration conditions and was involved in the ligations of $^{15}$N-labeled ERβ as well as FAM-ERβ-LBDs. Lidia Nieto performed the MD calculations and analysis for the manuscript and helped with the ERβ-NMR. Sonja A. Dames led and completed the ERβ-LBD NMR and the analysis of the NMR-data. Ralph Bosmans and Chan Vinh Lam provided access to the QTOF-LSMC under the supervision of Joost van Dongen. Sonja A. Dames and Lidia Nieto are thanked for the fruitful collaboration, described in this chapter and for contributions and support in the discussion of the results and conclusions drawn from these.

This chapter is based on a manuscript, written together with Lidia Nieto, Sonja A. Dames and Luc Brunsveld.
2.6 References

Ligand-independent activation of estrogen receptor by a single PTM


3 Phosphorylation-induced interaction between the estrogen receptor and Src-kinase

Abstract. Next to the direct involvement in the control and regulation of its target gene transcription, the estrogen receptor makes part of different cellular signaling pathways, which are regulated by post-translational modifications. The estrogen receptor tyrosine phosphorylation site, located in the C-terminal region of the receptor’s ligand binding domain (chapter 2) is embedded in a sequence motif, which closely resembles the high-affinity binding motif for SH2 domains in terms of electrostatic and physical character. SH2 domains are phospho-tyrosine recognition domains and often make part of the modular architecture of proteins involved in cellular signal transduction pathways, such as in protein tyrosine kinases. In fact, cellular studies in prostate and mammary cancer cells have indicated that the estrogen receptor forms a complex with the cSrc-tyrosine kinase. This interaction depends on the presence of the SH2 domain and the C-terminal ER LBD-region. The exact molecular determinants of this interaction have however not been characterized yet. This chapter summarizes the biophysical characterization of the cSrc-kinase SH2 domain with different (non-) phosphorylated ERα and ERβ truncation constructs, amongst which semi-synthetically phosphorylated ER LBD constructs, described in chapter 2.
3.1 Introduction

The ER is often described as ligand-activated transcription factor. Most research on the structure, function and regulation of ER has focused on its crucial role in the transmission of hormone signals towards target gene expression. In this classical and simplified model of ER activation, the accommodation of hormone in the receptors ligand binding pocket results in structural reorganizations which reveal the receptors nuclear localization sequence and thereby anchor the receptor in the nucleus, resulting in the binding of ER-dimers do their DNA recognition sequences. Harboring an agonistic hormone in its LBD, ER will associate with nuclear receptor coactivators which will lead to the assembly of the transcription machinery at that gene transcription locus, followed by target gene transcription and its specific downstream cellular effects. ER is, however, not only involved in the transmission of hormone signals to target gene activation. It is also involved in other signal transduction pathways. Moreover, ER is not only regulated by the available hormone ligands, but also by a variety of PTMs, such as amino acid side chain phosphorylation, acetylation or sumoylation.

Every cell harbors a plethora of signal transmission and communication pathways, which are largely regulated by the transmission of PTMs. One of the key modifications in these communication networks is tyrosine phosphorylation, which is catalyzed by protein tyrosine kinases (PTKs). PTKs can be transmembrane receptor proteins like the epidermal growth factor receptor (EGFR) or the insulin receptor. These receptor tyrosine kinases (RTKs) are activated by the binding event of their ligands such as EGF or insulin and translate this extracellular signal to an intracellular signal by activating a downstream signaling cascade which will eventually lead to a cellular response on the originally extracellular signal, such as the up-regulation of the expression of proteins involved in cell growths or metabolic regulation. The transmission of the extracellular signal to the downstream cascade often also occurs via phosphorylation of cytosolic protein targets such as the components of the cytoskeleton or other non-receptor tyrosine kinases (NRTKs), which in turn may also conduct the signal by phosphorylation of the specific tyrosines of their downstream protein targets. The phosphorylation of a tyrosine side chain can directly lead to the activation of a protein via phosphorylation-induced structural changes or it creates an interaction surface and thereby induces the interaction with other downstream members of the signaling cascade.
Phosphorylation-Induced Interaction Between the Estrogen Receptor and Src-Kinase

Figure 3.1 Schematic overview of a tyrosine kinase signaling cascade. Ligand (L) binding (e.g. EGF) to the cell-surface receptor (R) domain results in receptor (e.g. EGFR) dimerization and (auto-) phosphorylation (P) by the receptors tyrosine kinase domain (RTK). Phospho-tyrosines provide a binding surface for phosphotyrosine recognition domains (SH2), which thereby recruit the proteins, they are embedded in (e.g. a non-RTK; NRTK), resulting in tyrosine phosphorylation of the recruited protein (e.g. Src kinase). NRTK phosphorylation induces a conformational rearrangement, NRTK activation (*) and the subsequent phosphorylation of downstream target proteins (T; e.g. estrogen receptor), which induces an altered (activated) state of the target protein and thus a specific cellular response to the initially external signal.

The human PTKs are a large group of 90 closely related genes of which many are conserved amongst metazoans. The fraction of these PTKs which has been identified as NRTK counts 32 proteins, amongst which the largest subfamily of Src kinases. The members of the Src kinase family are involved in the regulation of several signaling processes, such as cytoskeleton reorganization and the immediate as well as adaptive immune response by T- and B- cell activation. Furthermore Src kinases play a role in pathological processes, amongst which different types of cancer. Src kinases are expressed in a wide variety of different cell types in multicellular organisms, facilitating intercellular communication through phosphorylation of amongst others adhesion-, growth factor-, and antigen- receptors. Src kinases share a conserved PTK architecture, built up from classical modular signaling protein domains (Figure 3.2).
Figure 3.2 (A) Schematic overview of the Src kinase family domain organization. With the N-terminal myristoylation for membrane localization of the kinase (curled line), the unique domain for varying function between different protein family members, the SH3 domain for the interaction with the proline rich regions (PXXP) of target proteins, and SH2 domain for the interaction with target proteins by recognition of their phosphorylated tyrosine residue, followed by the catalytic domain of the kinase, which specifically phosphorylates tyrosines of target proteins. Src kinase carries two tyrosine phosphorylation sites, which have either activating (A) properties or inhibitory properties (I) by intramolecular interaction with its own SH2 domain.

Src kinase structural organization has been investigated intensively by protein crystallography as well as protein NMR. The individual domains of the several Src kinases have been crystallized but also the full-length crystal structure of several Src kinases has been solved. Src kinases are targeted to the cell membrane via their N-terminal post-translational lipid-modification. The overall closely related Src kinases show the most inhomogeneity in their N-terminal 50-70 amino acid stretch (“unique” domains, Figure 3.2). C-terminally of this variable region, Src kinases build up from a conserved sequence of domains, being the SH3, SH2 and kinase domains. The C-terminal tail of the Src kinases contains a tyrosine phosphorylation site which auto-inhibits the kinase function in its phosphorylated state, by binding to its own SH2 domain and thereby locking the protein in an inactive conformation (Figure 3.2, “I”). Non-catalytic SH3 and SH2 domains are highly conserved protein domains, which specifically recognize proline rich stretches (PXXP) or phosphorylated tyrosines (pY) in interaction partners, respectively. These domains are thus facilitating the specific interaction with other proteins. Such a modular construction of the kinase proteins is a well described phenomenon in a large variety of proteins involved in signal transduction. The different combinations of conserved and structurally relatively independent protein domains in combination with few divergent domains determines the specificity of the multidomain proteins. The first of these modular domains which has been described and characterized in detail was the SH2
domain after its discovery in the 1980s in Src and other kinases. Early research on oncogenes, containing an SH2 domain, but no kinase domain, revealed that SH2 was involved in the interaction with various tyrosine phosphorylated proteins and indicated that it had a regulating effect on tyrosine phosphorylation, rather than a direct catalytic function. To date, there are about 120 SH2 domains identified in humans and many of them have been characterized by in vivo as well as in vitro approaches. Already from the first crystal structures of the cSrc kinase viral homologue vSrc kinase SH2 domain in complex with phospho-peptide a common binding mode for SH2 domains and their recognition sequences has emerged.

The selectivity and specificity of SH2 domain binding, meaning SH2 binding to their target sequences in only the phosphorylated state, has been investigated in great detail. The phospho-peptide binding site of SH2 domains can generally be divided in a phosphotyrosine recognition pocket and a ‘specificity determining region’ (SDR), which interacts with the amino acid residues located C-terminal of the phosphotyrosine. The phosphotyrosine recognition site acts as a switch, depending on the phosphorylation state of the target sequence by selectively establishing interactions only with the phosphorylated tyrosine but not with the non-phosphorylated tyrosine. A few specific amino acid residues in the phosphotyrosine recognition pocket have been identified to play a central role. The conserved R178 (numbering is based on cSrc kinase sequence for convenience) establishes direct ionic interactions with the phospho-group and is essential for effective protein interaction (Figure 3.3 right). All other interactions have been observed in several systems, but are less conserved than the interaction with R178. R159 and K206 have been reported to form a clamp around the phospho-tyrosine and S180 as well as T182 are hydrogen bond donors for the phospho-group. Furthermore H204 and S188 are also involved in the interactions for phosphotyrosine recognition. The residues involved in the SDR binding pocket (Figure 3.3 left) can establish long ranged interactions with the residues, located at the +1 and +2 position in the target protein with respect to the phosphotyrosine (K203 and R208, respectively). In addition, a number of residues were identified to form a hydrophobic binding pocket, which results in stabilizing contributions in the protein-protein interaction by a hydrophobic residue in the target sequence at position +3 with respect to the phosphotyrosine (Figure 3.3 left).

On top of this discriminating interaction, the SDR additionally modulates the strength of the interaction, which - in the competitive setting of the cytosolic environment
with many potential interaction partners around it, determines the specificity of the protein-protein interaction. At the same time it allows for cross-talk between different cellular signaling pathways due to the lack of stringent target binding specificity of SH2 domains. This fine balance between specificity and promiscuity in combination with the orchestrated effect of the often multimodular interactions, results in signaling networks rather than isolated downhill signaling pathways.\textsuperscript{56,57} This model is supported by the observation that SH2 domains bind phosphotyrosine-sequences with high affinity but also fast dissociation and rapid exchange.\textsuperscript{58}

![Cartoon representation of the vSrc SH2 domain](image)

**Figure 3.3** Middle: Cartoon representation of the vSrc SH2 domain (gray) bound to a phosphopeptide (black backbone and color coded by element). The two opposite interaction sites of the protein on either side of the central beta sheet (arrows) are shown overall in lighter gray (left of the dotted line, specificity determining region) and in darker gray (right of the dotted line, phosphotyrosine recognition site). The details of these interaction surfaces are highlighted in the enlargements on either side. The amino acid side chains, which are directly involved in the interaction with the phosphopeptide are highlighted in cyan and color-coded by element. The protein structure is adapted from 1SPS (vSrc kinase SH2 domain). For simplicity, the numbering of the residues refers to the sequence of the cSrc kinase SH2 domain.

SH2 domains share a general architecture of a central anti-parallel β-sheet located in between two α-helices (Figure 3.3).\textsuperscript{59} The phosphotyrosine undergoes direct ionic interactions with a conserved arginine residue within the SH2 domain\textsuperscript{53,54} and is embedded in a network of ionic interactions as well as aromatic interactions with its phenol ring.\textsuperscript{49} The binding site recognizing the phosphotyrosine is very similar between different SH2 domains.
Phosphorylation-Induced Interaction Between the Estrogen Receptor and Src-Kinase

domains, yet when taking the SDR site into account, differing overall binding modes can be recognized. For the Src kinase SH2 domains interaction with their target binding motifs, the so called ‘two-pronged plug-two-holed socket’ model has been described (Figure 3.3). This model summarizes the observation that for the effective binding to the Src kinase SH2 domains, two interactions are crucial for high affinity binding. In addition to the phosphotyrosine interaction, the interaction of a hydrophobic side chain at the +3 position with respect to the phosphotyrosine contributes to the binding affinity. The high affinity binding core sequence pYEEI, which originates from the hamster polyomavirus middle-sized tumor antigen, establishes such a ‘two-pronged plug-two-holed socket’ with its +3 positioned isoleucine. The two negatively charged residues at positions +1 and +2 additionally establish stabilizing ionic interactions with the SDR. Most K_D values for Src SH2 peptide binders determined by different biophysical techniques range between 0.1 and 1 µM. The extension of the core binding sequence revealed no large effects on binding affinities, which is supported by the observation that the four residues of the high affinity binding motif are the only residues which undergo direct interactions with the Src SH2 domains. The variation of the size of the hydrophobic side chain at the +3 position in the SH2 binding motif as well as the variation in atomic space for the accommodation of these side chains in the SDR in different SH2 domains accounts in part for the slightly varied SH2 binding modes and along with that the differences between the SH2 binding sequences. Differences in binding modes between SH2 domains are however often not large, as single point mutations are enough to change specificity of one binding mode to another.

The comparison between the identified pYEEI- motif for Src SH2 domains and less strong binders revealed that variations in the amino acid sequence immediately C-terminal of the phosphotyrosine result in significant yet tolerable changes in binding affinities. This observation lead to the recognition that specificity of SH2 protein-protein interactions is not solely derived from the direct discrimination between potential SH2 interaction motifs but also from the concerted interactions of the different modular protein domains as often found in signaling proteins.

The ER is involved in the regulation of many non-genomic pathways which affect processes such as cell proliferation, neuroprotection and vasorelaxation. The integration of both ER subtypes in cytosolic signaling cascades via the physical interaction with cSrc kinase has been demonstrated in mammary cancer cells with ERα and prostate cancer cells for ERβ. The interaction between cSrc kinase and both ER subtypes occurs in a
phosphorylation-dependent manner, indicating that the interaction via the Src kinase SH2 domain is involved. In fact, co-immunoprecipitation studies confirmed that the interaction between the two proteins depend on the presence of the cSrc SH2 domain and could be inhibited by phospho-peptide analogues of the C-terminal moiety of the ER, harboring the tyrosine phosphorylation site Y537 or Y488 in ERα and ERβ respectively (compare chapter 2). The clinical relevance of this interaction is given by the observation that agonistic steroid ligands of the ER as well as the androgen receptor (AR) induce the formation of a ternary complex between the AR, either subtype of the ER and the cSrc kinase resulting in downstream activation of the Src/Raf-1/Erks-pathway and as a consequence cell proliferation in prostate cancer cells, which can be antagonized by antagonists for either of the involved nuclear receptors specifically in cancer cells. The interaction between ERα and Src kinase can furthermore be stabilized ligand-independently by the modulator of non-genomic activity of ER (MNAR). MNAR interacts directly with ERα via its LXXLL motifs and with the Src kinase SH3 domain via its N-terminal PXXP motif. This protein interplay induces the activation of the Src/MAPK-pathway, which results in ERα phosphorylation on serine 118 and thereby enhanced ERα transcriptional activity. Also in this ternary complex, pY537 has been proposed as interaction surface for cSrc kinase and ERα. This latter Src kinase and ERα concerted signaling provides an example for the integration of ER function as cellular signaling protein and its function as nuclear transcription factor and indicates that these two functions should be considered when aiming for a better understanding of ER function.

The various non-transcriptional functions of ER, which have been established to date include qualitative cellular determination of the interaction regions of cSrc kinase and ER. Yet, a quantitative characterization of this interaction has not been reported before. The amino acid sequence of the proposed SH2 binding motif, located C-terminal of the pY537 and pY488 (pYDLLL) closely resembles the described SH2 high affinity binding motif (pYEEI). The quantitative and structural characterization at atomic resolution, however, is envisioned to provide direct proof and atomic detail of the protein-protein interaction. Therefore this chapter describes the in vitro characterization of the protein-protein interaction by biophysical methods such as fluorescence polarization and protein NMR. The semi-synthetic introduction of the tyrosine phosphorylation to the full-length ER ligand binding domain (LBD) allows the direct characterization of the phosphorylation-dependent interaction between the two protein domains.
3.2 Results and discussion

3.2.1 Expression and purification of cSrc and vSrc SH2-domains

The expression and purification of the homologous human (c) and viral (v) Src kinase SH2 domains, has been described at high yields of pure protein before.\textsuperscript{23,26} The reported protocol for protein purification was followed with minor adjustments. Both protein domains were expressed from pET-3a expression vectors in \textit{E. coli} cells designated for protein overexpression and subsequently purified by ion exchange chromatography (IEC) in the absence of a purification tag. The high degree of structural homology between cSrc and vSrc SH2 domain is reflected in a 97% sequence homology. vSrc SH2 domain contains three point mutations with respect to cSrc SH2 domain, being A168P, N224S and T246N. Due to these minor differences in the protein structure and amino acid sequence, the expression and protein purification was expected to be very comparable for both protein domains. Following 3 hours of standard IPTG-induced protein expression at 37 °C in \textit{E. coli BL21 (DE)} cells and subsequent one-step purification by a linear NaCl-gradient on sulfoethyl functionalized sepharose cation exchange column material, pure protein samples at a yield of about 30 mg per liter expression medium were typically obtained of both protein domains, eluting at a NaCl concentration of about 200 mM (Figure 3.4 and Figure 3.5).

\textbf{Figure 3.4} SDS PAGE analysis of the cSrc SH2 domain purification procedure. With samples from before induction of protein expression (1), taken at the time of induction of protein expression (2), after 3 hours of protein expression (3), unfiltered cell extract after cell opening and separation from cell debris (4), filtered cell extract after cell opening and separation from cell debris (5), pellet of cell debris after cell opening (6), flow through of the IEC material (7), different fractions of the protein peak, eluted from the IEC (lanes 8-13), sample of the concentrated protein pool (14) and Bio-Rad double color pre-stained protein marker in kD (15).
The SDS-PAGE analysis of the different steps taken during the expression and protein purification procedure confirms the high protein purity after single step purification of the SH2 protein domain (Figure 3.4, lanes 8-14). Samples taken after cell opening reveal that all SH2 domain protein is present in the soluble fraction and none in the insoluble (Figure 3.4, lane 5 and 6, respectively). Furthermore an overloading of the column material with the desired protein domain can be observed by the high abundance of SH2 domain in the flow through of the IEC column (Figure 3.4, lane 7). This overloading might be beneficial for the protein purity resulting from the one step purification. However, it is also well possible that the yield of the protein purification can be largely increased without compromising the purity of the protein by enlarging the ratio between the column binding capacity and the amount of SH2-enriched cell-extract. The same observations were made during the purification of vSrc SH2 domain.

![Graph](image)

**Figure 3.5** QTOF LC- and MS- spectra including the deconvoluted observed protein mass of purified cSrc SH2 domain ((A) right, expected mass 12286.9 Da) and (B) vSrc SH2 domain (right, expected mass 12298.9 Da) purification. The m/z spectra (A and B, right) of the indicated peak region (A and B, left; underscored and asterisk) of the LC-spectra were used for the deconvolution of the observed masses.
QTOF LCMS confirms the high purity of the two homologous protein domains and their sequence integrity after one-step purification (Figure 3.5). An additional purification step by size exclusion chromatography was performed with the intention to yield extra pure vSrc SH2 domain for the crystallization studies. However, no increase in purity could be detected (data not shown).

3.2.2 Characterization of the ER-Src kinase interaction via fluorescence polarization

Quantification of cSrc and vSrc SH2 domain interaction with ERα and ERβ H12 peptide analogues.

The interaction between cSrc-kinase and ERα as well as ERβ via the cSrc kinase SH2 domain has been demonstrated by pull-down assays from cancer cells. However, their direct interaction has not been quantified to date. A fluorescence polarization (FP) based assay allowed for the determination of the protein-protein interaction, employing fluorescent phosphorylated and non-phosphorylated (p)ER H12-peptide analogues (ERα: H-C K N V V P L (p)Y D L L L E M L D A H R L K(FAM)-CONH2 and ERβ: H-C K N V V P V (p)Y D L L L E M L N A H V L K(FAM)-CONH2, SH2 binding motifs underlined) of both ER subtypes. The homologous vSrc SH2 domain was included in the quantification of the protein-peptide interaction in order to detect potential effects on the interaction affinities, induced by the three differing amino acid residues with respect to cSrc SH2 domain. Serial dilutions of the SH2 protein domains at a constant concentration of fluorescently (FAM) labeled (p)ER FAM H12-peptides, allowed for the determination of their binding affinities (Figure 3.6 A and B). For comparison and as positive control, a peptide sequence (FAM-(p)BB; sequence: FAM-G(p)YEEI), based on the high affinity binding sequence has been included in the titrations (Figure 3.6 C).
Figure 3.6 FP titration of cSrc and vSrc SH2 domain to FAM-labeled phosphorylated (p) and non-phosphorylated ER H12 peptides. Concentration of the fluorescent peptides was kept constant at 50 nM.

The binding data in Figure 3.6 were analyzed by employing equation \( Y = B_{\text{max}} \frac{X}{K_D + X} + B_G \), with the maximum binding at saturating titrant concentrations (\( B_{\text{max}} \)) and the background binding at low concentrations of titrant (BG), yielding the \( K_D \) values listed in Table 3.1. For all tested peptide sequences, the phosphorylation of the peptide turns out to be an absolute necessity for the interaction with the SH2 domains, since no binding has been observed between non-phosphorylated sequences and the SH2 domains (Figure 3.6, Table 3.1). This is in line with the described switch function of tyrosine phosphorylation for the SH2 interaction with its target proteins.

Table 3.1 \( K_D \) values in \( \mu \text{M} \) determined by an FP binding assay (Figure 3.6), between fluorescently labeled (non-)phosphorylated peptides and vSrc and cSrc SH2 domain.

<table>
<thead>
<tr>
<th></th>
<th>ER(\alpha)</th>
<th>pER(\alpha)</th>
<th>ER(\beta)</th>
<th>pER(\beta)</th>
<th>FAM-BB</th>
<th>FAM-pBB</th>
</tr>
</thead>
<tbody>
<tr>
<td>cSrc SH2</td>
<td>no binding</td>
<td>11 ± 0.6</td>
<td>no binding</td>
<td>8.5 ± 0.5</td>
<td>no binding</td>
<td>0.45 ± 0.03</td>
</tr>
<tr>
<td>vSrc SH2</td>
<td>no binding</td>
<td>4.2 ± 0.2</td>
<td>no binding</td>
<td>5.7 ± 0.4</td>
<td>no binding</td>
<td>0.98 ± 0.05</td>
</tr>
</tbody>
</table>
The $K_D$ values, describing the interaction of cSrc and vSrc SH2 domain with the control-peptides FAM-(p)BB in the high nanomolar range (Table 3.1) are well in agreement with the binding affinity of the reported reference sequence. The affinities in the low micromolar range, revealed for the binding of Src SH2 domains to the pER-sequences are well in agreement with commonly observed affinities of SH2 domains for their target proteins. This confirms the relevance of this protein-protein interaction and suggests that also in the case of pER binding to the Src kinase SH2 domain, the multidomain interaction such as in combination with MNAR, may play an important role in the fine tuning of pER-kinase interactions by further stabilizing the protein-protein interaction via multivalency stabilizing effects. With respect to the high-affinity for SH2 binding motif pYEEI, the observed binding behavior for the pER-sequence pYDLL implies that the overall effect of the earlier described small individual effects of differences in the pER binding motif with respect to the high affinity binding motif add up to an overall lower interaction affinity. Interestingly, vSrc SH2 domain appears to bind slightly better to both pER-FAM-H12 peptides than cSrc SH2. The replacement of the isoleucine residue at the +3 position of the high affinity binding SH2 binding motif pYEEI by a leucine residue, as is the case in the ER-sequences (pYDLL), has been studied before and was not expected to disturb the high-affinity ‘two-pronged plug-two-holed socket’ binding mode, yet a small decrease in binding affinity could be expected from the leucine in this position. The aspartic acid side chain in the +1 position of the ER-embedded SH2 binding motif may have also contributed to the reduced binding affinity with respect to the glutamic acid present in the high affinity binding sequence, since it is expected to enlarge the already long-reached ionic interactions of the carboxylic acid group with the SH2 domain binding interface. Furthermore the hydrophobic leucine residue at position +2 in the pER-sequences is expected to disturb the long-reached ionic interactions, which the glutamic acid would undergo in the high affinity binder sequence at this position. However, since the effect of +1 and +2 position mutations have been reported to often be only marginal, the binding affinity between SH2 and the pER-sequences may be expected to be in the commonly observed range from the high nanomolar to the low micromolar regime.

In addition to the necessity of tyrosine phosphorylation for the interaction with SH2 domains, the sequence motif, into which the phosphotyrosine is embedded has already been shown to be of importance for the specificity of the interaction with SH2 domains. In order to identify the influence of a more extended binding sequence in contrast to a short interaction...
motif, different lengths of pERα- and pERβ- truncations were investigated by a FP based competition assay. Short pERα/β-sequences (pERα/β short, Figure 3.7 left and Figure 3.8 A-D), harboring the previously identified minimal SH2 interaction, were included in the competition assay next to the more extended H12-peptides (pERα/β long, Figure 3.7 middle and Figure 3.8 A-D) and the full-length pERα/β LBDs (Figure 3.7 right and Figure 3.8 E). Non-phosphorylated ERα/β long sequences were included in the competition assay as a negative control.

**Figure 3.7** Overview of the six different pER truncations, included in the dissection of the molecular determinants of pER-SH2 interaction. Left: pERα and pERβ short peptides; middle: pERα and pERβ long peptides; right: pERα/β LBD (adapted from PDB 3OLL).
Figure 3.8 Normalized polarization values (Norm. Pol.) of the FP competition assay of long (black) versus short (gray) phosphorylated and non-phosphorylated (squares and circles, respectively) ER peptides. (p)ERα-peptides (A and C) and (p)ERβ-peptides (B and D) as well as pERα and pERβ LBD (E) are titrated as competitors. SH2 domain (cSrc SH2 in A, B and E; vSrc SH2 in C and D) was present at a constant concentration of 4 µM and fluorescent reference peptide FAM-pERβ at 25 nM.
Table 3.2 IC$_{50}$ values in µM obtained from FP competition assay (Figure 3.8) including a 95 % confidence interval (CI) in brackets, quantifying the interactions from c/vSrc SH2 domains with long and short ER$\alpha$- and ER$\beta$- phospho-peptides.

<table>
<thead>
<tr>
<th></th>
<th>ER$\alpha$ long</th>
<th>pER$\alpha$ long</th>
<th>ER$\alpha$ short</th>
<th>pER$\alpha$ short</th>
<th>ER$\beta$ long</th>
<th>pER$\beta$ long</th>
<th>pER$\beta$ short</th>
<th>pER$\beta$ LBD</th>
<th>LBD</th>
</tr>
</thead>
<tbody>
<tr>
<td>cSrc SH2</td>
<td>n.b.$^a$</td>
<td>34 (30-38)</td>
<td>n.b.$^a$</td>
<td>18 (16-22)</td>
<td>18 (16-20)</td>
<td>n.d.$^b$</td>
<td>n.d.$^b$</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>vSrc SH2$^c$</td>
<td>n.b.$^a$</td>
<td>18 (7-11)</td>
<td>9 (7-11)</td>
<td>17 (13-22)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

$^a$ n.b. = no binding, $^b$ n.d. = not determined, $^c$ titrations were performed as singlets.

The IC$_{50}$ values, determined for the different pER truncation sequences by FP competition show a very similar competition behavior for most tested sequences (Figure 3.8 and Table 3.2). The differences in the IC$_{50}$ values for pER$\alpha$ long and pER$\alpha$ short in interaction with cSrc SH2 may hint at an additional contribution of the longer peptide sequence to the stability of the protein-peptide interaction. This observation is in line with the lower IC$_{50}$ value observed for the interaction between pER$\beta$ long and the vSrc SH2. However, when taking the CI into account, only the interaction between cSrc SH2 and pER$\alpha$ short seems to be slightly weaker than the overall observed binding character. Similar to the KD values determined in the direct binding experiments (Figure 3.6), IC$_{50}$ values in the low micromolar range were observed for short as well as long pER-peptides. Within the accessible concentration range of the pER LBDs a similar competition behavior compared to the peptide analogues has been observed (Figure 3.6 E). Overall the extension of the pER truncations does not appear to yield significant value for the protein-protein interactions, yet due to the uncertainty of the determined IC$_{50}$ values, this hypothesis requires further elucidation.

In addition to the effect of the lengths of the phospho-peptide, the effect of the amino acid sequences located immediately N-and C-terminal of the phosphotyrosine was investigated (Figure 3.9), by varying the N- and C-terminal lengths of the SH2 binding motif.
Figure 3.9 FP competition assay with different N-terminal and C-terminal lengths of the SH2 binding motif. Serial dilutions of the different peptide competitors VVPLpY (pERα ΔC), VPLpYD (pERα 1) and PLpYDLL (pERα 3) were titrated against a N-terminally fluorescently labeled reference peptide (VPVpYD) at a constant concentration of 20 nM and vSrc SH2 domain at 10 μM.

The binding affinities of the SH2 domain to the pER-peptides is clearly affected by the changes in N- and C-terminal lengths of the SH2 binding motif, surrounding the phosphotyrosine (Figure 3.9). Due to the lack of competitor saturation, estimated IC₅₀ values have been determined for the competition by pERα 1 and pERα ΔC. These emphasize the trend of clearly higher peptide concentrations needed for effective competition for pERα 1 (estimated IC₅₀ 0.1 mM) and pERα ΔC (estimated IC₅₀ 0.5 mM) compared to the IC₅₀ of 31 μM (CI: 27-37 μM), determined for pERα 3. This gradual decrease in binding affinities from pERα 3 to pERα ΔC confirms that, the closer the phosphotyrosine is located towards the C-terminal end of the short peptides, the lower its competitive capacity, emphasizing the importance of the C-terminal peptide sequence. Overall, the data suggests either a stabilizing role of the C-terminal amino acid residues or a destabilizing role of the N-terminal residues, or a combination of the two effects. Taken together with the earlier observation that an extension of the ER-peptides N-terminal of the proline, located at the (-2) - position with respect to the phosphotyrosine, does not affect the binding affinities, suggests that indeed the change in binding affinities is caused mainly by the changes on the C-terminal end of the peptides. This assumption is also very well in line with what has been reported in literature before.⁴⁹,⁶⁰ The introduction of the C-terminal aspartic acid, which is not expected to have vast effects on the binding affinity, since it has only been described to undergo weak long-reached ionic interactions, already increases the binding affinity approximately 5 fold (Figure 3.9) with respect to the pER-sequence, lacking any C-terminal amino acids (Figure 3.9 pERα 1 versus pERα ΔC). The additional extension by the leucines at the +2 and +3
position increases the binding affinity more than 10 fold (Figure 3.9 pERα 3 versus pERα 1), most probably by completing the ‘two-pronged plug-two-holed socket’ binding interaction.⁴⁹

Assigning functional and mechanistic effects to individual post-translational modifications in various regulatory networks is challenging due to the temporary nature of the modifications and the fact that they often occur in combination with other modifications. Because of this, the characterization of the atomic effect of single modifications is not trivial. In many cases, point mutations have been introduced to proteins in order to study the effect of changes in the electrostatic properties of amino acid side chains.⁸¹ A tyrosine phosphorylation site for instance can then be mutated to a glutamic acid or aspartic acid side chain in order to mimic the physical and electrostatic properties of a phosphorylation at this position. This type of modelling of protein properties by mutagenesis has resulted in many useful insights. In order to evaluate the effect of different substitutions of the phosphotyrosine, the effects of ERα- and ERβ- peptides with phosphoserine or glutamic acid as substituent have been investigated (Figure 3.10).

**Figure 3.10** FP competition assay of ERα- (A) and ERβ- (B) peptides with phosphoserine (pS) and glutamic acid (E) as substituents for the phosphotyrosine (pY). non-phosphorylated ER-peptides (Y) are included as negative control. N-terminal fluorescently labeled reference peptide (VPVpYD) is added at a constant concentration of 20 nM in combination with vSrc SH2 domain at 10 µM.

The introduction of phosphotyrosine substituents revealed that SH2 binding is entirely disrupted (Figure 3.10). Neither the phosphoserine nor the glutamic acid side chain can establish an interaction of the peptides with the SH2 domain. This observation emphasizes the remarkable selectivity of the SH2 domain towards the specific physical characteristics of the tyrosine side chain. The lack of residual affinity upon tyrosine substitution supports the relevance of the reported aminoaromatic interactions of the SH2 arginine 158 (R155 in vSrc SH2) with the phosphotyrosine phenol ring.⁵² Furthermore the
replacement of a tyrosine residue by glutamic acid for the *in vivo* investigation of protein interactions with SH2 domains cannot be considered a reliable model system.

*Dissection of the influence of ligand and coactivator binding on the interaction between cSrc and vSrc SH2 domain and FAM-labeled ERβ LBD.*

The interaction between the pER LBD and the Src kinase SH2 domain has not been characterized in a direct *in vitro* setting before. The introduction of a fluorescent label to the C-terminus of pER LBD H12, allowed for the direct observation of change in fluorophore polarization upon addition of increasing levels of SH2 domains, accompanied by pER LBD-SH2 domain complex formation (chapter 2 for details, **Figure 3.11**). The inclusion of ER LBD also results in an increased complexity of the interaction, since the ER LBD conformation and subsequently its SH2 binding behavior may be largely affected by the presence or absence of ligand or coactivator.

![Graph A](image1.png)![Graph B](image2.png)![Graph C](image3.png)

**Figure 3.11** FP titration of cSrc SH2 domain (A) and vSrc SH2 domain (B) to 200 nM FAM-labeled pERβ LBD (A and B) and ERβ LBD (C). The titrations were performed in the absence and presence of 10 µM estradiol (E2, graph A) as well as in the absence and presence of saturating concentrations SRC1 box2 coactivator-peptide (LXXLL) in the presence 10 µM E2 (B).
The $K_D$ values, determined for the interaction between pER LBD and v or cSrc SH2 domains, reside in the same low micromolar regime as the $K_D$ values determined for the pER peptide binder analogues, reported earlier (Table 3.3).

**Table 3.3** $K_D$ values in µM determined by an FP binding assay (Figure 3.6), between fluorescently labeled (non-)phosphorylated ER$\beta$ LBD and vSrc or cSrc SH2 domain.

<table>
<thead>
<tr>
<th></th>
<th>cSrc SH2</th>
<th>vSrc SH2</th>
</tr>
</thead>
<tbody>
<tr>
<td>LXXLL</td>
<td>(-)</td>
<td>(-)</td>
</tr>
<tr>
<td>E2</td>
<td>(-)</td>
<td>(+)</td>
</tr>
<tr>
<td>FAM-ER$\beta$ LBD</td>
<td>n.d.(^a)</td>
<td>no binding</td>
</tr>
<tr>
<td>FAM-pER$\beta$ LBD</td>
<td>5.2 ± 1.8</td>
<td>12 ± 4</td>
</tr>
</tbody>
</table>

\(^a\) n.d. = not determined.

Furthermore, the addition of agonistic ligand E2, coactivator (LXXLL) or a combination of the two does not affect this protein-protein interaction significantly. The $K_D$ values determined for the interaction between cSrc SH2 domain and ligand-saturated FAM-pER$\beta$ LBD in the absence of coactivator might indicate a tendency towards a lower binding affinity (Table 3.3), yet the large deviation on the value and the evaluation of the binding curve by eye as well as the observed value for the same conditions with the vSrc SH2 domain, suggest that this difference is not significant (Table 3.3 and Figure 3.11). These observations confirm that the quantification of this protein-protein interaction is sufficiently represented by the SH2 interaction with a peptide fragment as short as 6 amino acids (Figure 3.8 and Figure 3.7). In fact, in this FP assay set-up, the determination of $K_D$ values with the peptide analogues of pER is the preferred approach to yield reliable parameters, since the change in polarization upon binding of the smaller 12.3 kD SH2 domain to the larger 27.6 kD FAM-ER LBD is small and the accuracy of the resulting binding curves consequently low. To avoid the high errors observed for this FP method, fluorescent labeling of the SH2 domain instead of the ER LBD has been explored. Unfortunately low labeling efficiencies of the SH2 domain as well as limitations in the accessible concentration range of the LBD, due to protein precipitation above 150 and 350 µM (for pER$\alpha$ and pER$\beta$, respectively; see also limitations in Figure 3.8 E) were encountered. An alternative read-out method by microscale thermophoresis of the interaction between fluorescently labeled pER LBD with the unlabeled SH2 domain has been tried out at various buffer and capillary conditions.\(^{82}\) Unfortunately, non-specific adsorption of the ER LBD to the sample capillaries could not be eliminated. The
immobilization of either protein domain in a surface plasmon resonance setting has also been explored unsuccessfully.

The direct proof of the interaction between the two protein domains by FP (Figure 3.11 and Table 3.3) and the characterization of the influence of coactivator and E2 on the protein interaction provide valuable information for the follow up structural investigation of the protein-protein interaction.

3.2.3 Co-crystallization of vSrc SH2 domain with phosphorylated pERβ LBD

Crystallization of vSrc SH2 has been described under several conditions in complex with different phosphotyrosine-peptides. The semi-synthetic introduction of the phosphotyrosine 488 to the ERβ LBD allows for the structural investigation of a complex between SH2 domain and the pER LBD. Based on the reasonably high affinity character of SH2-pER complex formation, as revealed by the described FP assays and based on the fact that vSrc SH2 domain as well as pERβ LBD have successfully been crystallized individually, the co-crystallization of these two protein domains was explored. For this purpose, pERβ LBD was produced by protein semi-synthesis and vSrc SH2 domain was purified as described earlier with an additional size exclusion chromatography step in order to yield extra pure SH2 domain to prevent interference of residual impurities with the formation of a uniform co-crystallization protein crystal packing.

Expression, purification and ligation of pERB LBD for the co-crystallization of pERB LBD with vSrc SH2 domain.

The phosphorylated ERβ LBD, which was applied in the protein co-crystallization with vSrc SH2 domain, has been crystallized individually before. Referring to the conditions, described in literature, the same amino acid sequence of the pERβ LBD (ERβD261-K488MD thioester, ligated to the H12-peptide C481-L500DD including pY488 (H2N-C K N V V V P Y D L L L E M L N A H V L D D-COOH)) has been used for the co-crystallization set-up as well as the protection of the surface-exposed cysteine residues. The latter was performed by incubating the crystallization construct overnight with iodoacetic acid at 4 °C. The close vicinity of cysteine 481 to the phosphotyrosine 488 binding site for the SH2 domain and a resulting steric or electrostatic hindrance for the protein-protein interaction caused by the side chain protection generated the motivation to also include non-protected pERβ LBD. Protein thioester and semi-synthetic protein ligation
product were produced with small improvements of the reported protocol.\textsuperscript{83} The improvements of the protocol included the reduction of the reaction time from 48 to a maximum of 16 hours as well as the use of a smaller yet sufficient threefold molar excess of H12 peptide over protein thioester in contrast to a previously reported 20 fold molar excess. Incubation times of 3 hours at 4 °C have even been observed to be sufficient in some cases, yet to ensure the reaction to go to completion, a longer reaction time was chosen for, since no loss in protein stability, i.e. no protein precipitation, was observed over the longer incubation times. Purity and integrity of the protein thioester and ligated proteins were confirmed by QTOF LCMS (Figure 3.13) and SDS-PAGE (Figure 3.12). The first purification of protein thioester by affinity column chromatography and on-column intein-cleavage resulted in typical yields of about 10 mg protein thioester (including impurities, Figure 3.12 lane 2) per liter TB expression medium.

**Figure 3.12** Semi-quantitative SDS-page gel of ERβ LBD thioester, specifically designated for protein crystallization, before and after ligation. The same amounts of each protein (before or after ligation) were applied to lanes 2 to 4. Lane 1: protein molecular weight marker in kD; lane 2: ERβ LBD thioester; lane 3: ligation mixture of ERβ LBD thioester with the phosphorylated H12-peptide; lane 4: combination of samples 2 and 3; lane 5: Protein ligation product after affinity chromatography (E2-sepharose) purification; lane 6: sample 5 at a 2.5 fold increased amount.

Analysis of the protein samples for the co-crystallization by QTOF LCMS confirmed the purity of the semi-synthetic proteins. Additional peaks as observed in both ligation products did not correspond to protein impurities or residual unligated peptide (Figure 3.13 B and C).
Figure 3.13 QTOF LCMS spectra of (A) ERβ LBD thioester (calculated mass 25181 Da), (B) cysteine protected pERβ LBD (calculated mass 27692 Da) and (C) non-protected-pERβ LBD (calculated mass 27634 Da). The m/z spectra (A-C, right) of the indicated peak region (A-C, left; underscored and asterisk) of the LC-spectra were used for the deconvolution of the observed masses.

Cysteine side chain-protected pERβ LBD was concentrated at a 1:1 molar ratio with vSrc SH2 domain to a concentration of 4.3 mg/ml of the protein complex. The protein complex was combined with the protein crystallization conditions, contained in the JCSG+ screen in a sitting-drop set-up. The crystallization screen was stored at 4 °C. One crystal has been observed in condition E10 after about 1 month of incubation time (Figure 3.14), which
was left for longer incubation in the hope for crystal growth. Unfortunately after a longer incubation period, the crystal disappeared. The non-protected protein batch was included in two crystal screens. No crystal growth was observed. In addition to the separate purification of both protein domains prior to the mixing of the domains, size exclusion chromatography (SEC) of a mixture of pERβ LBD with vSrc SH2 domain was performed under several conditions on analytical as well as preparative scale in the attempt to purify the protein complex prior to crystallization screening. Unfortunately the protein domains did not assemble the protein complex under the tested conditions, which is most probably due to the dilution effect of the SEC method in combination with low amounts of initial protein complex. All attempts for co-crystallization were performed at E2 saturation.

![Image of crystals](image.png)

**Figure 3.14** Left: Image of a crystal observed in condition E10 of the JCSG+ crystallization screen for suitable conditions of pERβ LBD crystallization in complex with vSrc SH2 domain. Right: Image of a crystal observed during the condition screening for the co-crystallization of vSrc SH2 domain with pERα- and pERβ- peptide analogues.

Furthermore, a large variety of conditions for vSrc SH2 domain co-crystallization with pERα- and pERβ- peptide analogues was screened. Initial crystal growth has been observed under several conditions. Optimization of these conditions as well as longer incubation times, however, did not yield sufficiently large, symmetrical and diffracting protein crystals for the continuation of the X-ray diffraction and subsequent structure determination (**Figure 3.14**).

### 3.2.4 Solution NMR of cSH2

As an alternative to protein crystallization, protein NMR can be a useful approach for the characterization of the molecular basis of protein-protein interaction. The SH2 domain is a relatively small protein domain of about 100 amino acid residues. Moreover it has proven
to be stable at high concentrations and over long periods of time, which makes it an optimal target for protein NMR measurements. In fact, SH2 domains have been extensively studied by solution NMR and the previously published NMR assignment of the cSrc SH2 domain would allow the study of the cSrc SH2 domain interaction in a straightforward manner.\textsuperscript{18} NMR was used to monitor the binding of four pER peptides to cSrc SH2 domain, through titration of these unlabeled peptides to \textsuperscript{15}N-enriched SH2 domain and monitoring of the chemical shift perturbations (CSP).

**Expression and purification of \textsuperscript{15}N-labeled cSrc SH2 domain**

While several examples of \textsuperscript{15}N-labeled expression and purification of SH2 domains from diverse origins are known from literature,\textsuperscript{18} detailed protocols for the preparation of this protein have not been published. The quality and quantity of alternative labeling methods has also not been highlighted in the mentioned publications. Due to this lack in methodological detail, different expression protocols have been explored at first. In order to perform NMR-spectroscopy on the full-length protein domains, the cSrc SH2 domain was expressed in \textsuperscript{15}N-M9 minimal medium,\textsuperscript{84} supplemented with \textsuperscript{15}N-enriched NH\textsubscript{4}Cl as sole nitrogen-source. The protein domain was purified with a yield of around 1 mg per liter expression medium by one-step ion-exchange chromatography, as described earlier.

![Figure 3.15](image.png)

**Figure 3.15** SDS-gel electrophoresis analysis of the \textsuperscript{14}N-labeled SH2 purification steps (A). With samples of the supernatant after cell opening (1), pellet of cell debris (2), flow through after ion exchange chromatography (3), chromatography product peak fraction (4), protein marker in kD (5, 7), reference of unlabeled cSrc-SH2 domain (12.3 Da) (6). B: QTOF-LC and MS spectrum of the final protein sample with the major peak at retention time 2.7 corresponding to the protein mass 12464±1 Da.
Chapter 3

The expected molecular weight of the $^{15}$N-labeled SH2 ($^{15}$N-cSH2) domain is 12441 Da, assuming 100% introduction of the $^{15}$N-isotope. The observed mass of 12464±1 Da (Figure 3.15, Right) most likely corresponds to a complexation of a sodium ion with the protein. In order to yield higher amounts of the target protein, the amount of transfected *E. coli* BL21(DE) cells was enhanced by growing the cells in rich TB medium, followed by a transfer to the M9 medium in which protein expression was induced. In order to quantify the effect of the growth phase in rich medium, transfected *E. coli* BL21(DE) cells were grown in 2 x 25 ml overnight cultures of LB medium. These pre-cultures were transferred in parallel to either 1 L M9 expression medium or 2 L TB growth medium. Protein expression in cells grown in M9 medium was induced with 0.5 mM IPTG at a cell density of $A_{600nm} = 0.5$ and incubated for protein expression over night at 37°C and 220 rpm. The cells, grown in TB medium were separated from the growth medium by centrifugation at 6000 rpm for 10 minutes at 4 °C and subsequently re-suspended in 1 L M9 medium. Cells were allowed to re-suspend properly by shaking at 220 rpm at 37 °C, yielding a high density expression culture at $A_{600nm} = 1.8$. Subsequently, overnight protein expression was induced by 0.5 mM IPTG. Cell density after overnight expression had increased to $A_{600nm} = 6$ for the TB-grown cells and $A_{600nm} = 2.8$ for the cells, which were grown in M9 medium prior to expression. After cell opening by homogenization, the cell extract, containing the protein of interest, was purified by ion exchange chromatography (IEC) yielding 1 mg/L of less pure (Figure 3.16, left) $^{15}$N-cSH2 from the M9-grown culture and 9 mg/ L M9 medium $^{15}$N-cSH2 of high purity (Figure 3.16, right) from the TB-medium grown culture. Judging by the intensities of overexpressed $^{15}$N-cSH2 domain with respect to total cell extract, the relative abundance of $^{15}$N-cSH2 domain is higher in the TB-grown batch than in the M9-grown batch (Figure 3.16, lanes 4). Slight overloading of the IEC column with the $^{15}$N-cSH2 domain as concluded from the band at the height of $^{15}$N-cSH2 domain in the flow through of the IEC (Figure 3.16, lanes 7), prevents undesired immobilization of other highly abundant proteins to the IEC material. Starting from higher relative abundance of the protein of interest with respect to total cell extract as well as intended overloading of column material with protein of interest are known attributes to successful purification of target proteins.
The yield and purity of the $^{15}$N-cSrc SH2 domain, expressed in cells, which were grown in TB medium before induction were significantly higher. Consequently all NMR spectra were recorded with the $^{15}$N-cSrc SH2 domain obtained from the TB medium grown cells.

**Titration of cSrc SH2 domain with different truncation lengths of pERα and pERβ peptides**

In order to investigate the molecular determinants of the interaction between pER and the cSrc SH2 domain, three different lengths of pERα- and pERβ- truncated sequences were titrated into a solution with $^{15}$N-labeled cSrc SH2 domain and studied by NMR (Figure 3.7). The short pER peptide analogues, included in the interaction study, span the core interaction motif C-terminal of the phospho-tyrosine. The long pER peptide analogues represent the 22 amino acid long sequence, corresponding to the loop 11- H12 sequence in the LBD-context (chapter 2 and Figure 3.7) and can yield insight into potential additional stabilizing contacts between pER and SH2. The interaction study with pER LBD was envisioned to yield direct proof of the complex of the two domains and its molecular character.
The ratio between cSrc SH2 domain and peptide, used in the NMR samples was determined by a titration of peptide to the protein domain. Under equimolar conditions, significant chemical shift perturbations in the protein $^1$H-$^{15}$N-HSQC spectra were observed (Figure 3.17, red). However, also peak broadening appeared, which hints at an equilibrium between free and bound conformers and therefore indicates that the protein is not fully saturated under these conditions. A threefold molar excess of peptide over protein domain resulted in overall peak sharpening with respect to the equimolar situation and in most cases only a small additional change in the chemical shifts (Figure 3.17, black), indicating that the binding equilibrium is shifted far towards the protein-peptide complex. Another increase of peptide concentration to a six fold molar excess did not result in further significant changes in chemical shift or peak sharpness (Figure 3.17, green). To ensure that the binding equilibrium is fully shifted towards the $^{15}$N-cSH2-pER complex, all described cSrc SH2 spectra were recorded at a 6 fold molar excess of the different phosphorylated binding partners.

Figure 3.17. Superimposition of a zoomed area from the $^1$H-$^{15}$N-HSQC spectra of $^{15}$N-cSH2 in the presence of increasing amounts of pERβ short peptide. Included are the spectra of $^{15}$N-cSH2 in the absence of pERβ short peptide (blue), at equimolar concentrations of $^{15}$N-cSH2 and pERβ short peptide (red), at a threefold excess of pERβ short peptide over $^{15}$N-cSH2 (black) and at a 6 fold molar excess of the peptide over protein (green).
$^{1}$H-$^{15}$N-HSQC spectra of $^{15}$N-cSH2 were recorded in the absence of interaction partner as well as in the presence of each pERα and pERβ peptide, individually (Figure 3.7, left and middle). In order to quantify the effect of the peptide binding partners on the backbone orientation of the individual residues of $^{15}$N-cSH2 and to identify the protein regions involved in this interaction, the backbone chemical shift perturbations were calculated for the entire protein sequence in each case (Figure 3.18). The overall changes in chemical shifts upon addition of phospho-peptides are remarkably similar (Figure 3.18). Several residues revealed outstandingly high chemical shift perturbations upon saturation with the different phospho-peptides.

**Figure 3.18** Backbone chemical shift perturbations (CSP) of cSrc SH2 domain in complex with peptides pERα short (A), pERα long (B), pERβ short (C) and pERβ long (D). The peak corresponding to E181 was not detected in the bound state and is thus not included in these CSP graphs. Based on PBD file 1HCS, the regions of different secondary structure characteristics are highlighted, with α-helical regions in light gray, β-sheets in dark gray, and loop as well as unstructured regions in white. Two cut-off values at y= 0.5 and y=0.2 are represented in the graphs by dotted black lines.

Given the overall highly similar effects of peptide binding observed in all individual NMR samples, the same cut-off values for a graphic representation of the most relevant
effects were chosen for all samples by eye (at a backbone CSP of 0.2 and 0.5 ppm, Figure 3.18). Comparing the residues, which are most affected in the different samples, e.g. residing above the cut-off values, the different peptide binders reveal only marginal differences (Figure 3.18). The residues strongly affected by the binding of the four tested peptide sequences are therefore summed up in one model (Figure 3.19 right), with the two most strongly affected residues highlighted in red and the moderately affected residues in blue.

**Figure 3.19** Comparison of the SH2 residues reported in literature to be involved in and affected by the interaction with high affinity binding motif pYEEI (left) and the backbone regions affected by the interaction with pER peptides (right). Left: For comparison a representation of the SH2 domain originated from crystallography studies. The highlighted residues (cyan) refer to the residues which were reported to be directly involved in the interaction with the high affinity binding motif pYEEI (adapted from PDB ID 1SPS). Right: The SH2 backbone regions most affected by the interaction with the pERα- and pERβ- peptides based on the observed CSPs, are highlighted in the model of the cSrc SH2 domain in red (CSPs above cut-off = 0.5 ppm) and blue (CSPs above cut-off = 0.2 ppm). The model is adapted from PDB file 1HCS. The phosphotyrosine of the original PDB file is included (black) for visual orientation.

Overall, the observed CSPs occur for residues which are located in immediate vicinity to the residues which have been reported to be involved in the protein-peptide interaction of the SH2 domain with the high affinity binder pYEEI (Figure 3.19). Furthermore, no additional regions of the protein domain appear to be affected. Even though
no direct conclusion of the type of side chain interactions can be drawn from the CSP information for the SH2-pER interaction it is remarkable that only the backbone regions in the direct vicinity of the identified interacting SH2 residues are observed to be affected by the SH2-pER interaction. This observation implies a very similar binding mode for the pER peptide as for the high affinity binder pYEEI. The two CSP values, identified above the cut-off of 0.5 ppm are corresponding to residues, located in loops of the protein (Figure 3.19, right highlighted in red). These are located in immediate proximity to residues, involved in direct interactions with the binding motif (Figure 3.19, left). CSP values above the cut-off of 0.2 ppm but below 0.5 ppm are mainly located in the more rigid and hence less affected secondary structure regions (Figure 3.18 and Figure 3.19 right, highlighted in blue).

Applying the same buffer conditions for the cSrc SH2 domain $^1$H-$^{15}$N-HSQC recordings, as reported in literature allowed for a straight forward assignment of the $^1$H-$^{15}$N-HSQC peaks to the individual amino acid residues of the protein domain. All samples were therefore prepared at pH 6.0 (20 mM MES buffer, pH 6.0 with 50 mM NaCl, 1 mM DTT and 10 % D$_2$O). Unfortunately, the pI of the pER LBDs resides around 6.5 and exposing the ER LBD to the NMR buffer conditions would therefore be expected to cause at least partial precipitation of the ER LBD due to the removal of all charges on the LBD surface, which reliably leads to the disruption of the protein solvation and consequently protein aggregation. Since the pH of the NMR buffer is slightly lower than the pI of the LBD, stability of the LBD in the NMR buffer was tested. Unfortunately, protein precipitation was observed as a consequence of the transition to the NMR buffer. As an alternative, the pH of the NMR-sample could be adjusted to a pH-range in which the ER LBDs are known to be stable. Such an adjustment of the NMR buffer conditions possesses the inherent risk of an overall rearrangement of the protein $^1$H-$^{15}$N-HSQC ‘fingerprint’, resulting in a loss of the peak assignment, due to pH-sensitive changes in the protein fold and dynamics. To investigate this alternative approach, a series of $^1$H-$^{15}$N-HSQC spectra were recorded at pH 7.5 (Figure 3.20) in the absence as well as in the presence of phosphopeptide. Unfortunately, the chemical shifts of individual residues as well as the overall ‘fingerprint’ of the protein spectrum differs from the chemical shift assignment in literature (Figure 3.20). Several peaks even disappear entirely under these modified conditions.
Figure 3.20 Overlay of the $^1$H-$^{15}$N-HSQC spectra of cSrc SH2 domain at pH 7.5 (red) and pH 6.0 (black).

A closer look at the residues, affected by the altered sample conditions, revealed that all peaks which disappeared from the $^1$H-$^{15}$N-HSQC spectrum in the absence of phosphopeptide are located in loops or unstructured regions of the protein domain, judging from the secondary structure, reported in literature.\textsuperscript{18} The changes in the peptide-bound situation are less dramatic than in the peptide-free samples, nevertheless the population of lost peaks or peaks, which lost their assignment, partially overlaps with residues, involved in the interaction with the pER-peptides. Optimization of the buffer conditions to achieve both, a suitable environment for the ER LBDs as well as a recovery of all peaks of interest, followed the assignment of the recovered $^1$H-$^{15}$N-HSQC peaks can most probably be done. However, since the differences between the individual pER-peptides was only marginal, no large additional changes in the interaction properties can be expected by the elongation of the pER interaction partner to the full-lengths LBD and has therefore not been pursued. In order to get a more global idea on the interaction between the two protein domains, such as their orientation towards each other, the protein complex was investigated by SAXS measurements. Due to the requirement of relatively high concentrations, protein solutions of the more stable pER$\beta$ LBD, rather than pER$\alpha$ LBD and the suspectedly more stable vSrc SH2 domain were included in the measurements. Unfortunately, the proteins aggregated under all tested conditions.
3.3 Conclusion

The interaction between different pERα- and pERβ- truncated sequences and the cSrc kinase SH2 domain as well as its homologue vSrc kinase SH2 domain was successfully quantified. For the identification of the atomic determinants of the protein-protein interaction, an optimized expression protocol for ¹⁵N-labeled cSrc SH2 has successfully been developed. Subsequent 2D solution NMR confirmed that the pERα and pERβ constructs interact with the cSrc SH2 domain with high similarity and that their binding mode can be concluded to be similar to the interaction between SH2 and the high affinity binding motif pYEEI. Furthermore the value of phospho-mimicking mutagenesis of the tyrosine of interest for the characterization of the ER-Src kinase interaction was evaluated in vitro.

Direct quantification of the binding affinities between the pER truncations, including the semi-synthetic full-length pER LBD, and the SH2 domains by fluorescence polarization assays as well as the indirect measure of their interaction by competition assays revealed low micromolar affinities for their interactions. This is in line with the reported affinities of other SH2 interaction partners ranging from high nanomolar to micromolar affinities. Furthermore, phosphorylation of the ER-sequences acts as a switch to regulate the protein-peptide interaction by an all-or-nothing switch mechanism.

A similar binding mode and molecular determinants were identified for the protein-peptide interaction between pER truncations and cSrc SH2 domain by FP assays and NMR, explaining the similarity in binding affinities, compared to previously reported interactions. The importance of the residues of the SH2 core interaction C-terminal of the phosphotyrosine of interest was confirmed and quantified. The importance of the involvement of the individual binding motif residues reflected in the CSP observed for SH2 backbone regions in direct proximity to earlier identified amino acid side chains involved in the SH2-target interaction. In the greater picture, these observations imply that cSrc recruits phosphorylated ERα and ERβ by their SH2 domain interaction. A discrimination between pERα and pERβ cannot be accomplished based on this interaction alone, since the binding mode as well as binding affinities are close to identical between the two protein subtypes. Additional mechanisms for subtype discriminations like subtype specific ternary complex formation with other regulatory proteins may apply.

The introduction of amino acid side chains, which mimic the physical and electrostatic properties of the phosphotyrosine of interest has proven to not be effective in the in vitro characterization of the ER-truncation sequences with the SH2 domain. This
observation suggests that such mutations are not suitable for an in vivo approach of the characterization of this protein-protein interaction either. The reported observations point out the importance of a further development of alternative methods such as semi-synthetic approaches for the introduction of post-translational modifications and their use in the direct characterization of the effect of PTMs on protein-protein interactions.

Taken together, the data described in this chapter provides physical proof for the interaction of ER and cSrc kinase and emphasizes its strict phosphorylation controlled character. Furthermore the extension of pER truncations up to the length of the semi-synthetic pER LBD does not reveal additional stabilizing effects on the protein-protein interaction, which makes the investigation of the protein-protein interaction by peptide analogues feasible. Even though the interaction with the pER LBD may very well be varied by their complex fold and sequence extent, the core interaction has been described to be of high affinity and non-selective between the two ER subtypes. The reported data furthermore suggests that ligand-induced ER-Src kinase complex formation, as has been reported from cellular studies, as well as its disruption by antagonistic ligands is not explained by direct ligand effects since no significant effects of E2 on SH2-binding affinity was detected here.

In the future it will be of interest to evaluate further phosphorylation controlled interactions of ER with other proteins, containing SH2 domains as interaction-modules. Our data suggests that a combination of cellular induction of ER phosphorylation in combination with an in vitro characterization should be preferred over the introduction of phosphotyrosine-mimicking mutations at the position of the tyrosine of interest.
### 3.4 Experimental section

If not mentioned otherwise, all chemicals and materials were used as purchased from the usual sources and without further purification.

**Expression vectors.** The DNA construct for the expression of the vSrc SH2 domain was kindly provided by Gabriel Waksman.\(^2^3\) The expression vector, coding for the cSrc SH2 domain was designed based on the example of the established vSrc expression system (produced by GenScript). In both cases, the protein coding sequences were inserted into the individual expression vectors (pET-3a) with the restriction sites NdeI and BamHI yielding the expression vectors pET3-a_vSrcSH2 and pET-3a_cSrcSH2. In addition, the DNA sequence coding for cSrc kinase was optimized for protein expression in *E. coli* cells. The protein coding sequence, yielded a tag-free expression of amino acids 144-249 (protein ID: P00524) in the case of the vSrc SH2 domain and the equivalent amino acid sequence of cSrc SH2 domain (amino acids 147-252, protein ID: P12931). Protein expression of the SH2 domain in *E. coli* in principle results in an additional N-terminal methionine, yet methionine cleavage has been observed in some expression batches.

\[
\begin{align*}
\text{v-src} & \quad \text{QAEEWFGRIFKRIEERLILN}^{144}\text{ENPRGTFLVRESETTKGAYC}^{249}\text{LSV}^{250}\text{DFDNKGLNVKHYKIRKL} \\
\text{c-src} & \quad \text{QAEEWFGRIFKRIEERLILN}^{147}\text{ENPRGTFLVRESETTKGAYC}^{252}\text{LSV}^{253}\text{DFDNKGLNVKHYKIRKL} \\
\text{v-src} & \quad \text{DSG}^{144}\text{GFYTSRTQF}^{248}\text{SSLQQ}^{249}\text{LVAYYSKHADGLCHR}^{251}\text{LT}^{252}\text{VC}^{253}\text{PT}^{254} \\
\text{c-src} & \quad \text{DSG}^{147}\text{GFYTSRTQF}^{251}\text{SSLQQ}^{252}\text{LVAYYSKHADGLCHR}^{254}\text{LT}^{255}\text{VC}^{256}\text{PT}^{257} \\
\end{align*}
\]

**Figure 3.21** Alignment of the amino acid sequences of the SH2 domains from cSrc and vSrc kinase (protein ID: P12931 and P00524 respectively). The differences in the amino acid sequences of the two protein domains are underlined.

**Protein expression SH2 domains.** All flasks, growth media and solutions used for bacterial expression cultures were either autoclaved for 20 minutes at 121 °C or sterilized by filtration. The expression and purification of the cSrc and vSrc SH2 domains was done as described in literature with some adjustments of the protocol.\(^2^3\) The individual expression vectors pET3-a_vSrcSH2 and pET-3a_cSrcSH2 were transformed into *E. coli BL21(DE)* by classical 30 s heat-shock transformation at 42 °C followed by 2 minute incubation on ice. Subsequently cells were allowed to recover for about two cycles of cell division in Luria-Bertani (LB) medium at 37 °C in the absence of ampicillin. Transformed cells were plated and grown over night on LB-agar plates in the presence of 100 μg/ml ampicillin at 37 °C. Pre-cultures of 50 ml LB medium supplemented with ampicilline 100 μg/ml were inoculated with single colonies of freshly transformed cells and incubated shaking overnight at 37 °C and 220 rpm. 1 L LB medium expression cultures freshly supplemented with 100 μg/ml ampicillin in in 2 L baffled Erlenmeyer flasks were inoculated with the entire volume of the overnight pre-cultures and subsequently grown until a cell density at A\(_{600nm}\) of 0.8 at 37 °C and 220 rpm. Protein expression was subsequently induced by the addition of isopropyl β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM. Cells were kept on 37 °C and 220 rpm shaking incubation for 3 hours to allow protein expression. Protein-expression was finalized by separation of the expression cells from their growth medium by centrifugation for 10 min at 7000 rpm and 4°C. Harvested cells pellets were snap-frozen in N\(_2\) (l) and store at -80 °C until further usage.
**Protein purification SH2 domains.** The entire purification of SH2 domains can be performed on room temperature, yet protein solutions and buffers were kept cooled on ice whenever possible. Previously harvested expression cells were thawed from -80 °C on room temperature. Cell lysis was performed by ice water-cooled homogenization (EmulsiFlex C3, Avestin) at a minimum of two homogenization cycles in lysis-buffer with 20 mM Hepes at pH 7.5, 10 mM EDTA, 5 mM MgCl₂ and freshly added 5 mM DTT, 200 μM PMSF and 1 μg/ml DNase. Subsequent centrifugation for the separation of cell debris at 20000 rpm and 4 °C for 30 minutes yields the cell extract with the soluble target protein in the supernatant. Two HiTrap SP FF were coupled in-line (column volume each 5 ml, Ge Healthcare) and pre-equilibrated in starting buffer according to the column manual and at 4 °C with starting buffer. Filtered cell-extract was flown over the ion exchange column material applying a peristaltic pump. After coupling the protein-loaded IEC columns to the automated FPLC system (AKTApurifier), unbound protein was washed off with starting buffer and target protein was subsequently eluted with a gradient from 0 to 0.5 M NaCl in 20 mM Hepes, pH 7.5, 5 mM EDTA and 5 mM DTT. The protein eluted at a salt concentration of about 0.2 M. Protein peak fractions were pooled and concentrated by Amicon® Ultra 3K centrifugal filter devices, snap-frozen in aliquots and stored at -80 °C in the elution buffer conditions. For protein crystallization an additional size exclusion chromatography step (HiLoad 26/60 Superdex 75 column, Ge Healthcare) in buffer 20 mM Hepes, pH 7.5, 5 mM EDTA, 100 mM NaCl 5 mM DTT was performed.

**Expression of 15N-enriched cSrc SH2 domains.** Protein expression for 15N-labeling was performed in 1 L M9 minimal medium⁸⁴ (48 mM Na₂HPO₄·1 H₂O, 22 mM KH₂PO₄, 8.6 mM NaCl, 18.6 mM NH₄Cl, 2 mM MgSO₄, 0.4% glucose, 2 mM CaCl₂) supplemented with 100 μg/ml ampicillin. During autoclaving of the M9 medium, salt precipitation was observed. Consequently the medium was not autoclaved prior to protein expression but filtered sterile. The plasmid pET-3a_cSrcSH2 was transformed into E. coli BL21(DE) cells and grown overnight at 37 °C and shaking at 220 rpm in 25 ml LB medium pre-cultures containing 100 μg/ml ampicillin. In the protocol, lacking the additional cell-growth step in terrific broth (TB) medium M9 minimal medium expression cultures were inoculated with the full pre-culture volume and grown at 37 °C and 220 rpm until an optical density of A₆₀₀nm = 0.5. Protein expression was induced by addition of 0.5 mM IPTG and allowed to proceed over night at 37 °C and 220 rpm. Following protein expression, cells were harvested by centrifugation at 7500 rpm and 4 °C, snap-frozen in N₂(l) and stored at -80°C until protein purification. In the case of TB-grown cells, a 25 ml overnight LB medium pre-culture with 100 μg/ml ampicillin was used to inoculate 2 L TB medium (24 g/L yeast extract, 12 g/L peptone, 4 ml/L glycerol) supplemented with separately autoclaved TB buffer (9.4 g/L K₂HPO₄, 2.2g/L KH₂PO₄). Cells were grown at 37 °C and shaking at 220 rpm up until an optical density of A₆₀₀nm ~ 0.9 and were subsequently separated from the medium by centrifugation at 6000 rpm for 10 minutes at 4 °C. Cells were transferred to M9 minimal medium and re-suspended in the culture shaking incubator at 37 °C and 220 rpm. Subsequently, IPTG was added to a final concentration of 0.5 mM and cells were incubated at 37 °C and 220 rpm over night for protein expression. Cells were harvested by centrifugation at 7500 rpm and 4 °C, snap-frozen in N₂(l) and stored at -80°C until protein purification.

**ERα and ERβ LBD protein semi-synthesis.** The protein semi-synthesis was performed as described in chapter 2. The sequences of ERα (K302-K529) and ERβ (L260-K480) were cloned into pTWIN1 vectors⁸³ (New England Biolabs) and expressed recombinantly in E. coli BL21(DE) cells (EMD Millipore) as fusion proteins with a C-terminal Mxe intein domain to yield the C-terminal thioester after thiol-induced intein
cleavage. A C-terminal chitin binding domain was used for affinity purification on chitin-functionalized resin (New England Biolabs). After 20 h expression at 15 °C in TB medium supplemented with 100µg/ml of ampicillin, cells were opened by homogenization (EmulsiFlex C3, Avestin) in Buffer A (20 mM Tris at pH 7.5, 500 mM NaCl, 1 mM EDTA, 0.1 mM TCEP) supplemented with 200 µM PMSF, 1 µg/ml DNAse I (EMD Millipore), and 1 mM MgCl2. Affinity purification on chitin beads at 4°C in Buffer A, was followed by on-column intein cleavage at pH 8.0-8.5 in Buffer A with 200 mM MESNA for 16 hours at 4°C. After elution in the same buffer, the cleavage step was repeated. Protein concentration was increased by centrifugation in Amicon® Ultra 10K filter devices and the protein thioester was stored in the elution buffer at -80°C until ligation to H12-peptides. Altogether, this procedure typically yields around 15 mg protein thioester per liter TB medium. Prior to the native chemical ligation reaction, the respective H12-peptide was incubated at 20°C under reducing conditions (2 mM TCEP) to ensure reduced cysteine side chains. For the ligation reactions, protein thioester was incubated with 3 equivalents of the respective H12-peptide overnight at 4 °C and pH 8.0 in Buffer A supplemented with 2 mM TCEP and 200 mM MESNA. Subsequent purification by means of size exclusion ensured the removal of unligated peptide.

**Peptide synthesis.** All peptides were synthesized on Rink amide resin by standard Fmoc- solid phase peptide synthesis (SPPS) either manually or by automated peptide synthesis or a combination of both for the efficient introduction of fluorescent groups and the modified tyrosine residue. NMP (N-Methyl-2-pyrrolidone) was used as solvent for the entire synthesis protocol. Deprotection of the N-terminal Fmoc-protection group was done in 20 % (v/v) piperidine in NMP. Fmoc-amino acid building blocks were diluted from 200 mM stock solutions to yield a final twofold molar excess of building block over the resin-substitution, reported by the producer. Coupling reagents HBTU or HCTU were used. Coupling reagents were included at equimolar concentrations to the amino acid building blocks and together with 4 equivalents of N,N-Diisopropylethylamine DIPEA. The coupling of each amino acid building block was incubated for 10 minutes. This coupling step was repeated a second time under the same conditions. As building blocks for the introduction of the phospho-tyrosine and phospho-serine residues, Fmoc-Tyr(PO(OBzl)OH)-OH and Fmoc-Ser(PO3BzlH)-OH were used, respectively. After each double coupling step, unreacted N-terminal amines were acetylated and the final peptide sequences were N-terminally capped by acetylation in pyridine : NMP : acetic anhydride (1:3:1: ratio). Peptides were cleaved from the resin by a 94:2:2:2 ratio of TFA : H2O : TIS : EDT (TFA, trifluoroacetic acid; TIS, triisopropylsilane; EDT, 1,2-ethanedithiol) and precipitated in cold diethyl ether during incubation at -20 °C for minimal 1 hour or until precipitation was observed. Successful synthesis of the peptides was confirmed by analytical LCMS (Table 3.4). High purity of all peptides was assured by purification via preparative reversed phase HPLC with mass detection in adequate gradients from water to acetonitrile in the presence of 0.1 % TFA. Peptides were stored freeze-dried at -20 °C. An Fmoc-Tyr(PO(OBzl)OH)-OH (EMD Millipore) building block was used for the introduction of the phosphorylated tyrosine residue. The fluorescently labeled H12-peptides (Figure 3.6 A and B) were synthesized on a Tenta Gel R RAM resin (Rapp Polymere). A Dde-protected Fmoc-lysine was introduced C-terminal to the native H12 sequence. The orthogonal deprotection of the lysine side chain, allowed for coupling of 5- or 5/6-carboxy-fluorescein (FAM) to the lysine side chain NH2-group in a repeated overnight reaction in the presence of 4 equivalents of FAM, 3.6 equivalents of HATU and 8 equivalents of DIPEA. All other fluorescent probes were synthesized by coupling
of either a FAM- or a FITC- building block to the N-terminal amino-group of the peptide sequence (Figure 3.6 C and Figure 3.9 - Figure 3.10, respectively).

Table 3.4 Summary of peptide names, sequences, theoretical mass in Da and observed mass in Da.

<table>
<thead>
<tr>
<th>Peptide name</th>
<th>Peptide Sequence</th>
<th>Calculated mass (Da)</th>
<th>Observed mass (Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERa FAM-H12</td>
<td>H-CKNVPLYDLLLLMDAHLRK(FAM)-CONH2</td>
<td>2841</td>
<td>2841</td>
</tr>
<tr>
<td>pERa FAM-H12</td>
<td>H-CKNVPLYDLLLLMDAHLRK(FAM)-CONH2</td>
<td>2921</td>
<td>2921</td>
</tr>
<tr>
<td>FAM-BB</td>
<td>FAM-GYEE</td>
<td>967</td>
<td>967</td>
</tr>
<tr>
<td>FAM-pBB</td>
<td>FAM-GyYEE</td>
<td>1047</td>
<td>1047</td>
</tr>
<tr>
<td>pERa short</td>
<td>PLpYD</td>
<td>812</td>
<td>813</td>
</tr>
<tr>
<td>pERβ short</td>
<td>PVpYD</td>
<td>798</td>
<td>799</td>
</tr>
<tr>
<td>pERa long</td>
<td>H-CKNVPLYDLLLLMDAHLRKHA-COOH</td>
<td>2564</td>
<td>2564*</td>
</tr>
<tr>
<td>pERβ long</td>
<td>H-CKNVPLYDLLLLMDNAHLRG-COOH</td>
<td>2576</td>
<td>2576</td>
</tr>
<tr>
<td>ERa long</td>
<td>H-CKNVPLYDLLLLMDAHLRKHA-COOH</td>
<td>2564</td>
<td>2565*</td>
</tr>
<tr>
<td>ERβ long</td>
<td>H-CKNVPLYDLLLLMDNAHLRG-COOH</td>
<td>2496</td>
<td>2496</td>
</tr>
<tr>
<td>pERa 1</td>
<td>VPLpYD</td>
<td>684</td>
<td>686</td>
</tr>
<tr>
<td>pERa 3</td>
<td>PLpYD</td>
<td>812</td>
<td>813</td>
</tr>
<tr>
<td>Fluorescent probe</td>
<td>FITC-VpYD</td>
<td>1160</td>
<td>1162</td>
</tr>
<tr>
<td>pS-ERa</td>
<td>PLpSD</td>
<td>737</td>
<td>736</td>
</tr>
<tr>
<td>E-ERa</td>
<td>PLEDL</td>
<td>699</td>
<td>698</td>
</tr>
<tr>
<td>pY-ERβ</td>
<td>PVpYD</td>
<td>799</td>
<td>798</td>
</tr>
<tr>
<td>pS-ERβ</td>
<td>PVpSD</td>
<td>723</td>
<td>722</td>
</tr>
</tbody>
</table>

*Produced by Thermo Scientific

Fluorescence polarization direct binding and competition assays. All peptides were dissolved in the MST-buffer at pH 7.4. FAM-peptide assay concentrations were kept constant at 50 nM and SH2 domains were serially diluted from high micromolar to low nanomolar concentration ranges. Fluorophore polarization was detected by a Tecan Safire monochromator microplate reader (Figure 3.9, Figure 3.10 and Figure 3.11) and a Tecan Infinite F 500 microplate reader (Figure 3.6 and Figure 3.8). The resulting binding curves of the direct interaction assays were analyzed by formula (1), yielding the $K_D$ values for the interactions.

$$Y = B_{max} \cdot \frac{X}{(K_D + X)} + B_G$$

with the maximum binding at saturating titrant concentrations ($B_{max}$) and the background binding at low concentrations of titrant ($B_G$).

The competition assay conditions (Figure 3.9 and Figure 3.10) were based on reported assays,79 with minor adjustments. Fluorescent probe was kept constant at 20 nM (or 25 nM in Figure 3.8) and vSrc SH2 domain at 10 µM (or 4 µM in Figure 3.8) in 50 mM Tris-HCl at pH 7.5 with 100 mM NaCl and 2 mM DTT. FP and competition assays in Figure 3.6, Figure 3.8 and Figure 3.11 were performed under MST-buffer conditions (50 mM Tris at pH 7.4, 150 mM NaCl, 10 mM MgCl$_2$ 0.05 % Tween-20, 0.5 mg/ml BSA, 1 mM TCEP) optimized for the inclusion of the microscale thermophoresis read-out method as well as to minimize peptide aggregation and to include ER LBDs while assuring protein stability.88 The binding curves of the competition experiments as displayed in Figure 3.8 until Figure 3.10 were analyzed with formula (3), yielding the $IC_{50}$ values for the competition by the individual peptide sequences.

$$Y = P_{BOTTOM} + \frac{(P_{TOP} - P_{BOTTOM})}{(1 + 10^{((X - \text{LogIC}_{50})})}$$

with the top and bottom plateau (P).

Protein crystallization. Co-crystallization of vSrc SH2 domain was screened manually as well as automated under a wide variety of conditions. Co-crystallization screening of full-length pERβ LBD was performed by automated pipetting (Mosquito® Crystal) at a protein complex concentration of 4.3 mg/ml in the condition screening JCSG+ (Qiagen) in the case of cysteine side chain protected pERβ LBD in a sitting drop
approach on MRC2-plates. Conditions were incubated at 4 °C. pERβ LBD without cysteine side chain protection was included in the PACT primer condition screen (Molecular Dimensions) and Wizard classic 1 & 2 condition screen (Emerald Bioscience) at a protein complex concentration of about 10 mg/ml.

**Solution NMR of 15N-cSrc SH2.** NMR samples of 15N c-SH2 protein were prepared in MES buffer (pH 6.0) to a final concentration of 150 μM, containing 10% D2O. 1H-15N HSQC (heteronuclear single quantum correlation) spectra were collected at 25 °C on a Bruker 750 MHz spectrometer. 1H-15N HSQC spectra were acquired with 1024 complex data points in the 1H dimension (t2) and 256 increments in the indirect dimension (t1). Data sets were zero-filled to 2048x1024 before Fourier transformation using Topspin (Bruker), and analyzed with CARA 1.8.4. The chemical shifts of vSrc-kinase SH2 domain, as reported by Xu et al., were used as reference for backbone assignment (Figure 3.22). Peaks were labeled with the one-letter code for amino acids and the sequence position number. The trials of 1H-15N HSQC spectra recorded at pH 7.5 were done in a 20 mM Hepes buffered solution.

The backbone chemical shift perturbations ($\Delta\delta_{av} = \left\{[(\Delta\delta_{HN})^2+(\Delta\delta_{N}/5)^2]\right\}^{1/2}$) were calculated for the entire protein sequence and for every investigated peptide interaction.

![Figure 3.22](Image)

**Figure 3.22** 1H-15N-HSQC of cSrc SH2 domain at pH 6.0. Including the assignment of the peaks. Amino acid residues and numbers are allocated to the respective peaks.
3.5 Acknowledgements

The protocol for the peptide synthesis of the fluorescently labeled ERb-H12-peptides was established by Christian Haase during the synthesis of the FAM-ERb-peptides and later applied by Mark Balk to the FAM-ERa-peptides. Lidia Nieto designed and performed the NMR-experiments at the national NMR facility (SONNMRLSF) in Utrecht with support by Hans Wienk and Rolf Boelens. Lidia Nieto furthermore determined the CSPs resulting from these experiments. Andrea Schmidt performed many crystallization screens of vSrc SH2 domain with ERα and ERβ-peptides that were synthesized by Mark Balk during his internship project. The pET3a-vector coding for the vSrc SH2 domain was a kind gift of Gabriel Waksman. Lech Milroy, Lidia Nieto and Marcel Scheepstra were involved in the guidance of the student project on the different SH2 binder peptide sequences. Ralph Bosmans and Chan Vinh Lam provided access to the QTOF-LSMC under the supervision of Joost van Dongen. Seppe Leysen helped with setting up the full length pERβ LBD-vSrc SH2 crystal screens.
3.6 References

(13) Published online: 20 November 2000; doi:10.1038/sj.onc.1203957 2000, 19.
(44) SH2 Domain https://sites.google.com/site/sh2domain/home (accessed May 26, 2014).
4 Estrogen receptor α hinge region modifications and their influence on the hCaM-interaction

Abstract. The estrogen receptor α hinge region mediates the concerted action of the N-terminal and C-terminal domains of the nuclear receptor. Post-translational modifications (PTMs) of the hinge region have been reported to have strong regulatory impact on receptor transactivation. In different types of breast cancer, the malfunction of hinge region PTM-regulation has been identified to result in enhanced estradiol-sensitivity and cell-proliferation as well as resistance to endocrine therapy. The molecular basis of the underlying regulatory mechanisms has however not been addressed in the past, which is in part due to the inaccessibility of the homogeneously modified protein samples. This chapter provides an approach to the dissection of the molecular determinants involved in intra- as well as in inter-molecular regulation or ERα activation by PTMs. The described mechanisms cover direct structural effects of PTMs and the more indirect effects of hinge region modifications on ERα activation, mediated by interaction with regulatory protein interaction partner human calmodulin. In order to investigate the effects of hinge region modifications on protein structure and dynamics in a protein-context, a semi-synthetic introduction of hinge region modifications into ERα protein constructs is envisioned. This chapter summarizes and discusses the efforts towards such semi-synthetic protein constructs.
4.1 Introduction

The ERα hinge region refers to a ~60 amino acid short non-conserved and presumably mainly unstructured sequence stretch, located C-terminally of the receptors DNA binding domain (DBD) and N-terminally of the ligand binding domain (LBD, Figure 4.1). The effects and function of individual hinge region modifications have been characterized in cellular assays. However, the assignment of specific effects of an individual PTM in high mechanistic detail is challenging in the cellular context due to low spatial resolution and potential uncontrolled cross-talk with other modifications. The work described in this chapter is therefore dedicated to the establishment of a molecular system by which a detailed molecular characterization of hinge region PTM effects on intrinsic hinge region structural character and protein-protein interactions can be investigated with biophysical methods.

ERα hinge region PTMs and their functional and clinical relevance

The effect of estrogens in their target tissues is mediated by the estrogen receptors ERα and ERβ. The individual cellular response to the hormone signal however depends on many parameters such as the cell type and state, e.g. by different promoters, expression levels of the subtypes ERα and ERβ as well as their isoforms and potential mutations therein. Next to ligand-activation or activity modulation, ER activity is highly regulated by PTMs, which have been studied most intensively for the ERα subtype. Individual ER modifications and the interplay between different modifications have been described to influence ER function as transcription factor as well as its function as a member of complex signaling networks. The classical activation mechanism of the transcription factor ER is triggered by ligand-binding and the resulting activating conformational rearrangements of the receptor, which is presumed to reveal the nuclear localization sequence located in the ERα hinge region. Regulation of ER activity by PTM can occur in a ligand-induced manner or ligand independently. The extensive regulation of ER by different covalent modifications including phosphorylation, acetylation, methylation, sumoylation and ubiquitinylation has diverse effects on receptor activity. Phosphorylation of serine 118, located in the AF-1 region of ERα has, for example, been identified to play a role in receptor dimerization and promotes ligand-dependent as well as ligand-independent activation of target gene expression. Also outside the AF-1 region, the ERα hinge region presents a high density of target residues for PTMs, amongst which hinge region-specific lysine acetylation,
sumoylation and ubiquitylation but also serine phosphorylation, which occurs throughout
the full-length of the ER sequence (Figure 4.1).

Figure 4.1 Schematic overview of a selection of ERα hinge region PTMs (not drawn to scale). With
the hinge region located C-terminal of the DNA binding domain (DBD) and N-terminal of the ligand
binding domain (LBD).

ERα hinge region modifications are correlated with an altered mediation of ligand
effects by ERα. Prevention of K303-acetylation or ubiquitylation by a K303R mutation
which often occurs in breast cancer tumors is associated with an increased proliferation rate
as response to E2. Furthermore phosphorylation on S305 has revealed strong correlation
with adverse effects of the anti-estrogen tamoxifen, mediated by conformational
rearrangement of ERα. These observations make the hinge region organization of
particular interest for in-depth molecular investigation. A dissection of the effects of hinge
region modification is also of strong clinical relevance, since about 70% of all breast cancers
are classified as ER positive, meaning that their growth is dependent on ER activity.
Therefore the two major approaches in anti-breast cancer treatment are the inactivation of the
ER by anti-estrogens like tamoxifen and the inhibition of estrogen production by aromatase
inhibitors. Unfortunately approximately 40% of the patients that undergo endocrine
therapy reveal intrinsic resistance to the treatment or develop an acquired resistance
throughout the treatment. Several mechanisms for intrinsic as well as acquired resistance to
anti-estrogens have been identified, such as overexpression of steroid receptor
coadaptors, increased activity of ER interacting partners AP-1 and NF-κB, ligand-
dependent ER signaling pathways, specific effects of the different ERα and ERβ
isoforms, and the interplay and abundance of certain ER PTMs. The complex molecular
mechanisms which lead to resistance against endocrine therapy in different breast cancer
patients will thus still require in-depth investigation. An important pragmatic approach in
order to achieve advance in individualized breast cancer treatment is to identify biomolecular markers which can be correlated to endocrine response and thus provide an indication for the most promising treatment of individuals.\textsuperscript{7,27,35} The phosphorylation status of ER\textsubscript{α} hinge region S305 (Figure 4.1) is considered such a promising biomarker.\textsuperscript{5,7} The phosphorylation of S305 is catalyzed by PKA and associated with enhanced receptor transcriptional activity,\textsuperscript{5,6,36,37} which is supported by the observation that phospho-serine mimicking S305E mutation, anchors the receptor dimer to transcription sites of target genes and is sufficient for the activation of certain target genes ligand-independently.\textsuperscript{5,36–38} Up-regulated PKA activity, like it has been observed in tamoxifen resistant breast cancer cells, and thus S305 hyperphosphorylation can result in increased agonistic effects of tamoxifen by inducing an altered interaction mode with ER coactivators.\textsuperscript{5,6} An interplay between S305 phosphorylation and the acetylation of its neighbored lysine K303 has been implied by the inhibition of lysine acetylation by the phospho-serine (pS) mimicking S305E mutation.\textsuperscript{3,39} Other acetylation sites in the direct vicinity of the latter are K302 and K299. All three lysine side chains are acetylated by p300. Just like K299 and K302-acetylation, modification of K303 by p300 has an overall suppressing effect on ER\textsubscript{α} transcriptional activity.\textsuperscript{39} This effect may be the result of a combination of the inhibition of S305 phosphorylation and thereby inhibition of the activating effect of pS305 on ER\textsubscript{α} transcriptional activity and the interference with ER\textsubscript{α}-activating protein-protein interactions such as the dimerization promoting interaction with human calmodulin (hCaM).\textsuperscript{40} However, the molecular details of the influence of hinge region modifications on ER\textsubscript{α} protein-protein interactions are not studied in much molecular detail yet, due to the lack of control over homogeneous introduction of individual modifications in a cellular context. ER activity is furthermore modulated indirectly by calcium signaling regulated phosphatases and kinases and via direct interactions with calcium binding messenger hCaM and the ER\textsubscript{α} hinge region.\textsuperscript{40–42}

In order to study the influence of hinge region PTMs on ER\textsubscript{α} structure and function as well as on its protein-protein interactions, truncated synthetic sequence analogues of the differently modified hinge region sequence can be used for initial characterizations. Yet, ultimately a semisynthetic protein construct of the modified N-terminal hinge region sequence and the C-terminal ER\textsubscript{α} ligand binding domain, reacted via native chemical ligation (NCL, Figure 4.2), will provide a toolbox for the investigation of the PTMs in a protein domain context.
Figure 4.2 Protein semi-synthesis of ERα hinge region-LBD\textsubscript{S309C-L554} protein constructs from synthetically modified hinge region peptide thioesters (left) and recombinantly expressed ERα LBD (right) by native chemical ligation, yielding a library of differently modified protein constructs.

For the introduction of different hinge region modifications by NCL, the N-terminal peptide moiety needs to be synthesized with a C-terminal thioester (Figure 4.2).\textsuperscript{43–45} Furthermore the recombinantly expressed C-terminal protein moiety will need to yield an N-terminal reactive cysteine.\textsuperscript{46,47} The introduction of an N-terminal cysteine to a protein domain is not trivial, since regular protein expression naturally induces an N-terminal methionine to the protein sequence. Depending on the C-terminal amino acid sequence of the protein of interest, the N-terminal methionine may also be cleaved off at sufficient yield within the host cell, to high extend.\textsuperscript{48,49} Furthermore, several techniques have been developed to circumvent the N-terminal native methionine for the purification of recombinantly expressed proteins, displaying N-terminal cysteines.\textsuperscript{48,50–54}

It has been reported specifically for the ER subtype α,\textsuperscript{55} that the contributions of ERα ligand-independent N-terminal (activation function 1, AF-1) and its ligand-activated C-terminal activation function 2 (AF2) to the receptor activation are concerted by the ERα hinge region.\textsuperscript{56} The mechanism of individual effects of hinge region modifications as well as combinations of different modifications on the intrinsic hinge region structure are therefore of interest and will be explored in this chapter. Next to enzymes, catalyzing the PTM of the hinge region and the removal of the modifications,\textsuperscript{4,39} other regulatory proteins such as calmodulin have been reported to bind the ERα hinge region.\textsuperscript{40,42,57,58} In this chapter, the
characterization of the interaction between ERα hinge region and hCaM and the influence of different modifications hereon is described.

In order to dissect the molecular effects of ERα hinge region modifications on ligand binding, cofactor interaction and other protein interactions e.g. with hCaM, the introduction of these modifications into the ERα LBD is envisioned. Ultimately, a semi-synthetic protein construct spanning DBD, specifically modified hinge region and the LBD can be envisioned. This will allow the investigation of the different hinge region modifications embedded between the two protein domains, which are presumably affected by these modifications in structure and function. Towards such an extended multistep semi-synthetic ERα protein construct, the introduction of hinge region modification to the hinge region in a semi-synthetic N-terminally extended LBD was approached as a first step. Several strategies for the construction of such a semi-synthetically modified hinge region - LBD protein construct are described in this chapter.
4.2 Results and discussion

4.2.1 Peptide analogues of the modified ERα hinge region

In order to investigate the individual effects of ERα hinge region PTMs on intrinsic hinge region structural properties as well as for the dissection of the interplay between the different modifications, 14 amino acid long peptide sequences of the highly modified C-terminal end of the hinge region (P295-L310, protein ID: P03372) were synthesized, decorated with various combinations of the different occurring PTMs (Table 4.1).

<table>
<thead>
<tr>
<th>Peptide Number</th>
<th>Peptide sequence</th>
<th>PTM</th>
<th>Calculated MW/ Da</th>
<th>Observed MW/ Da</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>PLMIKRSKKNSLAL</td>
<td>-</td>
<td>1639</td>
<td>1639</td>
</tr>
<tr>
<td>(2)</td>
<td>PLMIKRSK*KNSLAL</td>
<td>K303</td>
<td>1681</td>
<td>1681</td>
</tr>
<tr>
<td>(3)</td>
<td>PLMIKRSK*KNSLAL</td>
<td>K302</td>
<td>1681</td>
<td>1681</td>
</tr>
<tr>
<td>(4)</td>
<td>PLMIK*RSKKNSLAL</td>
<td>K299</td>
<td>1681</td>
<td>1681</td>
</tr>
<tr>
<td>(5)</td>
<td>PLMIKRSK*KNSLAL</td>
<td>K302, K303</td>
<td>1723</td>
<td>1723</td>
</tr>
<tr>
<td>(6)</td>
<td>PLMIK<em>RSK</em>KNSLAL</td>
<td>K299, K302, K303</td>
<td>1765</td>
<td>1765</td>
</tr>
<tr>
<td>(7)</td>
<td>PLMIKRRKKNS*LAL</td>
<td>S305</td>
<td>1719</td>
<td>1719</td>
</tr>
<tr>
<td>(8)*</td>
<td>PLMIKRSKK*SNSLAL</td>
<td>K303, S305</td>
<td>1761</td>
<td>1753</td>
</tr>
<tr>
<td>(9)*</td>
<td>PLMIK<em>RSK</em>KNS*LAL</td>
<td>K299, K302, K303, S305</td>
<td>1845</td>
<td>1845</td>
</tr>
</tbody>
</table>

*Lysines are acetylated and serines are phosphorylated. *Peptides are not included in CD spectroscopy.

Overall peptide synthesis was successful, however the yield of peptide (9) was extraordinarily low by manual as well as automated synthesis, most probably due to the less efficient introduction of the many modified amino acid building blocks. The mass difference detected for peptide (8) was observed in the approach by manual as well as automated synthesis. The origin of this mass difference could not be identified. Peptide (8) and (9) were therefore not included in the following characterization by circular dichroism (CD) spectroscopy.
Characterization of the PTM effects on the intrinsic hinge region structural character by CD spectroscopy

The potential effects of the individual PTMs, as introduced via peptide synthesis (Table 4.1) were investigated by CD spectroscopy. CD spectroscopy finds frequent application to yield valuable data on the conformational signature of proteins and polypeptides.\textsuperscript{59-62} Increasing the overall hydrophobicity of the peptide surroundings by adding 2,2,2-trifluoroethanol (TFE) is considered to mimic the dielectric environment of a protein-embedment by peptide backbone desolvation and thereby allow an extrapolation on the tendency of intrinsic structural characteristics of peptide sequences in a protein-context.\textsuperscript{63,64} The effect of a range of TFE concentrations was therefore explored (Figure 4.3 and Figure 4.4).

![Figure 4.3](image-url)

**Figure 4.3** Mean residue molar ellipticity (MRE) for peptide (1) (A) and peptide (7) (B) in the presence of different percentages (v/v) of TFE in sodium phosphate buffered aqueous solutions. Samples in the presence of 30% TFE are included in duplicate originating from individual peptide stock solutions (30% TFE_1 and 30% TFE_2, respectively).

In **Figure 4.3**, two representative summaries of the behavior of the individual peptides in different solvent backgrounds are presented. The error between the same samples from independent peptide stocks is negligible. In the absence of TFE and in the presence of 10% TFE, a random coil signature with a single minimum around 195 nm was observed. This observation is in line with the assumption that the hinge region has a low level of secondary structure. The increase of TFE to 30%, however, reveals an α-helical signature spectrum with a maximum at around 190 nm and minima at about 208 and 222 nm. This observation implies an intrinsic tendency of the hinge region peptides to adopt a helical conformation in an altered solvent surrounding. In the exemplified cases of peptide (1) and (7) in **Figure 4.3**, the helical signature increases even more upon the introduction of 50%
TFE. These overall tendencies have been observed for all tested peptides (Figure 4.4) with clear effects of the peptide modifications (Figure 4.4 and Figure 4.5). Since the spectra at 30 % TFE show the most pronounced differences between the differently modified peptides, the effects of the modifications are further on discussed based on the results obtained in the presence of 30 % TFE. (Figure 4.4, 30 % TFE).

![Graph showing mean residue molar ellipticity (MRE) for peptides (1) – (7) at different percentages (v/v) of TFE in sodium phosphate buffered aqueous solutions.](image)

**Figure 4.4** Mean residue molar ellipticity (MRE) for peptides (1) – (7) at different percentages (v/v) of TFE in sodium phosphate buffered aqueous solutions. Spectra are recorded under identical conditions apart from the TFE concentrations of 10 %, 30 %, 50 % and in the absence of TFE (0 %). The legend and axis titles for 30 % TFE apply to all graphs.

The helical character of unmodified peptide (1) clearly increases in the presence of 30 % TFE, compared to 10 % TFE (Figure 4.4 and Figure 4.3 A). However, the helix representative signature of all singly acetylated peptides (2)-(4) is more pronounced than for the unmodified peptide (1) (Figure 4.4). The signature of peptide (2) with acetylation of K303 appears to be most helical of all tested singly acetylated peptides. The introduction of K302-acetylation into peptide (2), yielding peptide (5) reveals a significant reduction of helical character, emphasizing the relevance of cross-talk between different modifications on hinge region structure and thus function. A combination of all three lysine acetylations by peptide (6) clearly induces a strong helical tendency to the sequence region (Figure 4.4). Altogether these observations reveal a stimulatory effect of lysine acetylations on the peptides α-helical character. Interestingly serine phosphorylation has a destabilizing effect on
the sequence α-helical character as indicated by a less pronounced helical signature of peptide (7) compared to all other peptides.

As a measure of the effect of different modifications on the helical CD-signature of the different peptides, the MRE-ratios at 222 versus 208 nm were calculated (Figure 4.5). The comparison of the MRE-ratios in the absence of TFE (Figure 4.5, white bars) and in the presence of 30 % TFE (Figure 4.5, gray bars) emphasizes the role of the peptide’s intrinsic helical tendency in a protein-mimicking solvent-surrounding. This observation is in line with the reported structural data of the hinge region sequence of interest in interaction with hCaM.

![Figure 4.5](image)

**Figure 4.5** Ratios of the mean residue ellipticity at 222 nm with respect to 208 nm. The ratios are plotted per peptide number (1) until (7) for the samples without TFE (0 % TFE, white bars) and at 30 % TFE (gray bars).

The MRE-ratios confirm the overall stabilizing effect of the lysine acetylations on the peptide’s helical signature (Figure 4.5 gray bars (2)-(6) versus (1)). The MRE-ratios furthermore reveal the different contributions of the individual lysine acetylations to this overall stabilizing effect. The introduction of K303-acetylation has the strongest effect on the MRE-ratio, by increasing it from 0.6 in the unmodified peptide (1) to 0.8 in peptide (2). The other two lysine acetylations, have a smaller, yet also positive effect on the MRE-ratio by increasing it to 0.7 (peptides (3) and (4)). The introduction of the less stabilizing K302-acetylation to peptide (2) even counteracts the stabilizing effect of K303-acetylation as measured by a decrease in MRE-ratio to 0.7. The combination of the latter two modifications with the third tested acetylation or K299, however, results in a recovery of the high MRE-ratio of 0.8, equivalent to the single K303-acetylation of peptide (2). The contrasting effect of serine phosphorylation in peptide (7) as observed in Figure 4.4 is confirmed by the
quantification of the decrease in MRE-ratios from 0.6 to 0.5, when comparing the unmodified peptide (1) with the serine-phosphorylated peptide (7), respectively (Figure 4.5). All of the observed tendencies are reflected in the MRE-ratios for the samples in the absence of TFE (Figure 4.5, white bars) with the exception of peptide (6), which shows a less pronounced minimum at 195 nm and a shift of that minimum towards higher wavelengths, actually indicating a tendency towards more helicity. This observation also marks the limits of the reliability of the MRE-ratio analysis and will thus not be used as base for interpretations. The observed helical character stabilizing effects of lysine acetylation and destabilizing effects of S305-phosphorylation were confirmed by NMR measurements. Also the switch from unstructured character to helix-formation upon transition to the 30 % TFE-environment was confirmed by NMR (data not shown). Next to the cross-talk between the different modifications as described above, it is also of interest, to which extent the different modification patterns of the individual peptides contribute to or counteract the effect of the solvent-exchange. This effect can be expressed in the change of the MRE-ratios from the TFE-free samples compared to the samples containing 30 % TFE (Figure 4.5, white versus gray bars). The absolute increase in MRE-ratios upon solvent-exchange is very similar and resides around 0.5 for the unmodified peptide (1) and peptides (3)-(5). In comparison, the stabilizing effect of K303-acetylation is additionally emphasized by the larger increase by 0.6 of the MRE-ratio from 0 to 30 % TFE. Also the earlier described stabilizing effect of triple modification of peptide (6) is underlined by a larger increase by 0.7 of the MRE-ratio upon solvent exchange from 0 to 30 % TFE. Also upon solvent exchange, peptide (6) revealed the lowest increase of MRE-ratio by 0.4. Since this increase is smaller than the increase for unmodified peptide (1), this comparison again confirms the destabilizing effect of serine phosphorylation. Taken together, the observations on the helical character of the modified ERα hinge region sequences confirm that the peptide sequences, originating from the supposedly naturally unstructured protein domain, indeed show no helical character when exposed to an aqueous environment. However, the results reveal a significant potential to stabilize intrinsic helicity by change of their environment and the peptide modifications were revealed to differentially affect this tendency.

In order to elucidate the molecular picture of the described effects, a helical wheel projection of the ERα hinge region peptide sequences was added to the analysis (Figure 4.6). This wheel projection points out that an α-helix in this sequence region would display all its positively charged amino acids on one face of the helix (Figure 4.6 left, highlighted in red),
whereas the surface on the opposite of the helix exposes hydrophobic and polar side chains. This model would imply a relative destabilization of the helix in the unmodified state (Figure 4.6 left) by the repulsion of the positively charged residues. Consequently the neutralization of at least part of the charges by means of lysine acetylation can be expected to increase helix-stability (Figure 4.6 right). This interpretation of the sequence projection is in line with the observed effects in the CD experiments (Figure 4.4 and Figure 4.5). According to the helical wheel projection, K299 and K303 are located in close proximity, implying potential repulsions between the residues in the unmodified state. This implication of the helical wheel projection is in line with the observation that acetylation of both residues have a positive effect on the peptide’s helical signature (Figure 4.4 and Figure 4.5). In its triple acetylated form, the polarity of the helix-surface decreases significantly and it would lose its amphiphilic character (Figure 4.6 right), which appears to be of positive influence on the helical signature (Figure 4.5) and could also be of significance for the interaction with other proteins.

![Helical wheel projection of peptide sequences](image)

**Figure 4.6** Helical wheel projection\(^{67}\) of the analyzed peptide sequences with hydrophilic (blue), hydrophobic (yellow) and positively charged (red) amino acid side chains. Displayed are the unmodified peptide (1) (left) and triple acetylated peptide (6) (right). The individual lysines, which are acetylated in peptides (2)-(4) are indicated by their peptide numbers (right).

In addition to exerting stabilizing effects on secondary structure, the lysine side chain modifications may regulate protein-protein-interactions directly via modulating their ionic interfaces. This mechanism of regulation strongly suggests itself for the interaction between the ER\(\alpha\) hinge region and hCaM. The ER\(\alpha\) hinge region lysines 303 and 299 have been shown to make part of the protein-protein-interface. In fact, the side chains of these two
lysines are oriented in close proximity to a glutamic acid side chain of hCaM (Figure 4.7). This binding mode is highly similar for the N- and C-terminal lobe of hCaM.\textsuperscript{40}

![Image](image.png)

**Figure 4.7** N-terminal lobe of hCaM (gray) interacting with the hinge region of ER\(\alpha\) (black) via direct interactions between ER\(\alpha\) lysine 299 (K299) and K303 with the glutamic acid side chain of hCaM (E12). The picture is made from PDB file 2LLO.\textsuperscript{40}

Since the molecular basis for interaction between hCaM relies on the interaction with ER\(\alpha\) lysine residues which are subject to post-translational modification, the regulation of this interaction by PTMs is of interest. The \textit{in vitro} characterization of the interaction between hCaM and differentially modified ER\(\alpha\) hinge region is therefore envisioned to yield insight in the interplay of ER\(\alpha\) regulation by hCaM and the potential control of their interaction via ER\(\alpha\) PTMs.

### 4.2.2 hCaM protein expression and purification

For the investigation of the ER\(\alpha\) hinge region interaction with hCaM by biophysical techniques \textit{in vitro}, 149 residue full-length hCaM was expressed and purified from transfected \textit{E. coli} BL21 (DE) protein expression cultures. The expression and purification protocol for human hCaM was based on protocols reported in literature and successfully applied with minor adjustments.\textsuperscript{68} For the direct interaction studies between modified ER\(\alpha\) hinge region peptide-analogues and hCaM via ITC and fluorescence polarization assays, a wild type and unlabeled hCaM was produced. Furthermore a study of the hCaM dynamics and the effect of ER\(\alpha\) hinge region interaction thereon by NMR was envisioned. For this purpose \textsuperscript{15}N-enriched S17C (hCaM_S17C) or A128C (hCaM_A128C) mutants were produced, intended for the covalent introduction of paramagnetic probes via a metal
chelating-tag. The cysteine mutations have been described as compatible with hCaM function and structure before.\(^{69}\) Consequently, the metal chelating tag can be linked covalently to the single cysteine side chains, which were introduced by mutagenesis into the otherwise naturally cysteine-free hCaM sequence.

Wild type hCaM expression and purification

Wild type (wt) hCaM was successfully expressed and purified. The purification protocol was inspired by protocols described in literature, expected to yield as much as 250 mg of pure protein per liter expression culture after eight hours expression time.\(^{68}\) Purification of hCaM following the described procedure, yet with a reduced expression time of 4 hours, resulted in only minor enrichment of the target protein and consequently protein purity was not satisfying for the intended in vitro characterization of ER\(\alpha\)-hCaM interaction (Figure 4.8 left, sample 3 versus right sample 8). An additional size exclusion purification step yielded approximately 30 mg of highly pure hCaM per liter expression culture (Figure 4.8 right, lane 13).

![Figure 4.8](image)

**Figure 4.8** SDS PAGE analysis of the hCaM purification with the reference Bio-Rad double color pre-stained protein marker in kD (1). Left: SDS PAGE of the hCaM purification steps with a sample of the expression culture before induction of protein expression (2) and after the expression time of 4 hours (3), flow through of the phenyl-functionalized sepharose column material (4) and subsequent wash fractions (5-7). Right: SDS PAGE of the additional size exclusion chromatography (SEC) step with a sample of the product fraction eluted from the phenyl-sepharose column (8) and the different peaks observed during SEC (9-13), amongst which the product peak (13).

The SDS PAGE analysis of the protein purification reveals that the target protein expression is enhanced by the addition of IPTG (Figure 4.8 left, lane 3 versus 2) and target protein hCaM appears to remain selectively on the column material (Figure 4.8 left, lane 4). However, non-target protein was not observed to be washed off the column material in the
following washing steps (Figure 4.8 left, lanes 5-7), which could be either due to the dilution of the samples by large volumes of washing buffer or it could indicate the unintended capture of non-target protein on the column material, resulting in poor purity of the target protein (Figure 4.8 right, lane 2). Subsequent SEC purification yields pure hCaM (Figure 4.8 right, lane 13). The purity and sequence integrity of hCaM was confirmed by QTOF LCMS analysis (Figure 4.9).

Figure 4.9 QTOF LC- and MS-spectra of the purified hCaM (expected mass: 16708 Da).

The limited enrichment of target protein after phenyl-sepharose column chromatography (Figure 4.8 right, lane 8) may be improved by increasing the ratio between target protein and total cell extract as well as by lowering the amount of column material with respect to the applied protein extract. Higher levels in protein overexpression may be achieved by elongating expression times to the described 8 hours rather than the performed 4 hours.68 Overnight expression times for the subsequently explored 15N-enriched hCaM expression will be described as a first approach to higher protein purity after one-step protein purification (vide infra).

15N-labeled hCaM_S17C and hCaM_A128C mutants

In order to investigate the changes in protein-protein interaction between hCaM and ERα hinge region upon hinge region acetylation and the behavior of hCaM dynamics as affected by this interaction on a residue level resolution, two hCaM cysteine mutants were recombinantly expressed and purified. To yield a high amount of pure and 15N-labeled hCaM, transformed E. coli BL21 (DE) cells were first grown in rich TB medium, following the protocol established for the production of 15N-cSrc SH2 domain in chapter 3.70 Following cell-growth in 1 L TB medium for each mutant, cells were separated from the growth medium and transferred into 1 L M9 medium by centrifugation and re-suspension when having reached a cell density at A600nm of about 1.71 Purity after single-step purification like
described for wt unlabeled hCaM was confirmed for both mutants by SDS PAGE (Figure 4.10).

![SDS PAGE analysis](image)

**Figure 4.10** SDS PAGE analysis of the hCaM purification steps of cysteine mutants hCaM_S17C (left) and hCaM_S128C (right) with the reference Bio-Rad double color pre-stained protein marker in kD (1). Protein expression levels were monitored before (2) and after (3) induction by IPTG. After cell-opening, the soluble (4) and insoluble (5) fraction were analyzed. During column chromatography, samples were taken of the column flow through (6), 2 subsequent washing steps (7 and 8) and of the elution fraction before (9) and after concentration (10).

SDS-PAGE analysis of the $^{15}$N-enriched hCaM cysteine mutants reveals significantly improved purity after one-step purification, compared to non-labeled hCaM wt (lanes 10 in Figure 4.10 and Figure 4.8 lane 8, respectively). Comparison of the ratios between target protein and total cell extract with the wt expression by gel image analysis allows an approximation of about a 10% increase of target protein with respect to the total cell extract in the case of the cysteine mutants. This increase may have contributed to the increased purity after one-step protein purification. An additional noteworthy difference between the purification procedures of wt hCaM and the hCaM mutants is the effective washing steps after protein application of the gravity column material (lanes 7 and 8 in Figure 4.10 versus lane 5-7 in Figure 4.8), which may at least partially be responsible for the higher purity of the hCaM mutants. A comparison of the samples taken before and after protein induction of protein expression confirms significant amounts of target protein expression prior to expression induction by IPTG (Figure 4.10 lanes 2 and 3, respectively), like has also been observed during wt-expression (Figure 4.8 lane 2). The main problem of a leaky expression system in the applied expression protocol is the fact that the introduction of $^{15}$N-enriched nitrogen to the target protein can only occur after the transition of the protein expression cells to the minimal M9 medium. In this way, a labeled as well as an unlabeled protein population is being produced. With regard to the influence of the inhomogeneous
nature of the protein sample, the main disadvantage of the unlabeled protein present is a potential dilution effect of the signal intensity at a given protein concentration. However, compensation of this effect could be achieved by increasing overall protein concentration. Also the amount of peptide binder that needs to be added will be higher in order to saturate all present binding sites. Altogether, the population of unlabeled hCaM should not interfere with the envisioned NMR studies as long as the labeled protein population is prominent enough to yield a reasonable signal to noise ratio within a reasonable recording time of the NMR-spectra. A quantification of the extent of $^{15}$N-incorporation can be approximated by QTOF-LCMS analysis (Figure 4.11 and Figure 4.12).

![Figure 4.11 QTOF LC- and MS-spectra of the purified $^{15}$N-enriched hCaM_S17C (A) with the m/z-spectrum of the major LC-peak (*). Deconvolution of the selected m/z range (B) yields two mass-populations (x). The lighter molecular weight population corresponds to the unlabeled protein (expected mass after N-terminal Met-cleavage: 16721 Da). The population of higher weight corresponds to a continuous distribution of differentially $^{15}$N-enriched hCaM_S17C with a peak at the most abundant mass of 16877 Da (expected mass at 82 % labeling: 16876 Da)

The analysis of the $^{15}$N-enriched hCaM_S17C mutant by LCMS-spectrometry confirms the premature target protein expression prior to induction of protein expression by IPTG in M9 medium (Figure 4.11, B and ‘x’ and Figure 4.10, samples 2). Based on the QTOF-MS detection, the labeled hCAM_S17C represents 95 % of the overall protein sample.
Furthermore a large continuous distribution of differentially $^{15}$N-enriched hCaM_S17C was observed (Figure 4.11, B and ‘x’) with its peak in abundance at the mass of 16877 Da. This mass corresponds to the theoretical mass of the protein population of which 82 % of all nitrogen atoms are occupied by the $^{15}$N-isotope instead of the $^{14}$N-isotope. Since the $^{15}$N-labeling is occurring randomly by the bacterial expression machinery and the intended NMR-experiments are performed in bulk, the overall signal of the $^{15}$N-labeled atoms is expected to be sufficient for the envisioned experiments.

Very similar observations as described for the hCaM_S17C mutant were made for the hCaM_A128C mutant (Figure 4.12).

**Figure 4.12** QTOF LC- and MS-spectra of the purified $^{15}$N-enriched hCaM_A128C (A) with the m/z-spectrum of the major LC-peak (*). Deconvolution of the selected m/z range (B) yields two mass-populations (x). The lighter protein population corresponds to the unlabeled protein (expected mass after N-terminal Met-cleavage: 16738 Da). The population of higher weight corresponds to a continuous distribution of differentially $^{15}$N-enriched hCaM_A128C. The most abundant mass of 16890 Da equals the mass expected for 81 % labeling (16890 Da).

In the case of hCaM_A128C mutant, the percentage of labeled protein is 90 % of the total protein sample and the peak of the population of labeled protein resides at 16890 Da, which corresponds to 81 % random introduction of the $^{15}$N-isotope throughout the most
abundant protein population based on the abundance, detected by the QTOF ion-count (Figure 4.12). Also in this case, the labeling efficiency is expected to be sufficient for a good signal to noise quality of the envisioned NMR-recordings.

### 4.2.3 Biophysical characterization of the interaction between hCaM and ERα hinge region

As has been reported in literature, hCaM contains two binding sites for ERα hinge region. The interaction between hCaM and the peptide analogue of the binding site has been characterized by protein NMR and a $K_D$ of about 1 µM has been reported. The influence of ERα hinge region PTMs on this interaction however has not been addressed before.

Hinge region peptides spanning the ERα hinge region residues R287-L308 in analogy to the sequences reported in literature, were synthesized including the three lysine acetylations (K299, K302 and K303) individually and in combination. The binding affinities of these peptides to hCaM wt were investigated in direct biophysical affinity determinations by isothermal calorimetry (data not shown). The affinity of the unmodified peptide was determined to be around 3 µM at a binding stoichiometry of 1:2 hCaM to ERα-peptide, which is in agreement with the reported interaction characteristics. In summary, the determined $K_D$-values for the interaction between hCaM and the acetylated ERα hinge region peptides reveal that all acetylations lower the affinity of hCaM for the binding site. When comparing to the binding-affinity of the non-modified peptide, the acetylation of K299 and K303, which are both involved in ionic interactions at the protein-protein-interface (Figure 4.7), reduces the binding affinities 22- and 13-fold, respectively. The acetylation of K302 has a measurable, yet smaller effect of a 6-fold decrease on the binding affinity. Also the introduction of the K302 acetylation in addition to the K303-acetylation, which already lowered the binding affinity on its own, does not have a large additional effect on binding affinity towards hCaM. These observations clearly confirm the hypothesized influence of lysine acetylation on the ERα hinge region interaction with hCaM and identify K303 and K299 acetylation as the key residues for the regulation of their binding affinities by PTMs.

In order to interpret the atomic effects of hinge region acetylation on hCaM-binding, the protein-peptide interaction was characterized by protein solution NMR. A challenge in the investigation of hCaM by NMR is the pseudo-symmetry of its two binding sites. The high resemblance in the N- and C-terminal lobes of hCaM results in a relevant overlap of peaks in the protein fingerprint as obtained from $^1$H-$^{15}$N-HSQC-spectra. Probably due to this
substantial overlap, the only structural study of the hCaM-ERα complex reported thus far has been obtained for the separate N-terminal and C-terminal lobes of hCaM.\textsuperscript{40} This approach however, excludes the ability to gain information on the relative orientation between the N- and C-terminal lobes, their dynamics in the cooperative binding process and its significance in the binding of interaction partners. In the approach explored here, the pseudo-symmetry can be disturbed by the introduction of a paramagnetic metal ion in one of the lobes of hCaM, which results in an enhanced nuclear relaxation (paramagnetic relaxation enhancement, PRE) and pseudo contact shifts (PCS) for the residues in its conformational proximity over 40 Å and also in residual dipolar couplings (RDCs).\textsuperscript{73–78} Due to these paramagnetic effects, this strategy offers access to detailed long-range structural information, allowing the study of the individual binding sites and the relative orientation between the N- and C-terminal lobes within the hCaM-ERα complex.

For the paramagnetic studies of both, the N- and C-terminal lobes of hCaM, a metal chelating tag, inspired by earlier reported examples will be linked to the sulfhydryl functional group of the two mutated constructs hCaM_S17C and hCaM_A128C and saturated with diamagnetic and paramagnetic metal ions La\textsuperscript{3+} and Dy\textsuperscript{3+}, individually.\textsuperscript{79–81} To ensure, the proper introduction of the tag as well as the metals and the correct protein fold of the expressed protein mutants in the first place, protein fingerprints are recorded of the individual steps of modification (\textbf{Figure 4.13} and \textbf{Figure 4.14}).
Figure 4.13 (A) Comparison of the $^1$H-$^{15}$N-HSQC of $^{15}$N-enriched hCaM wild type (wt, black) with $^{15}$N-enriched hCaM_S17C (red). (B) Overlay of the un-tagged $^{15}$N-enriched hCaM_S17C (red) and the same mutant after introduction of the metal-chelator tag (green).

The comparison of the $^1$H-$^{15}$N-HSQC -spectra of the $^{15}$N-enriched hCaM_wt with the $^{15}$N-enriched hCaM_S17C reveals minor differences between the fingerprints of these two protein variants (Figure 4.13 A), that can mostly be accounted to the effect of the single mutation. The introduction of the chelator-tag was monitored by $^1$H-$^{15}$N-HSQC -spectra until no further changes were observed in the spectra. The final protein sample after removal of potentially unreacted tag (Figure 4.13 B, green) confirms the presence of one single tagged protein species, which indicates complete conversion of the reaction between protein and metal-probe. It furthermore reveals overall few and small chemical shift perturbations of NH-resonances in close proximity to the tag. These observations allow the assumption that the overall protein fold is stable and not significantly affected by the introduced modifications.

For the introduction of the metal ions, the chelator-tag was pre-saturated individually with the metal ions, prior to the reaction with the hCaM cysteine mutants. The subsequently recorded $^1$H-$^{15}$N-HSQC -spectra (Figure 4.14 A), containing a quantitative saturation of the protein-linked chelator-tag with the dia- and paramagnetic metal, revealed pronounced changes as result of significant pseudo contact shifts. Furthermore, an overlay of
the $^1\text{H}-^{15}\text{N}$-HSQC spectra, recorded with the hCaM_S17C linked to the paramagnetic metal-saturated tag in the absence and presence of un-modified ERα-peptide (Figure 4.14 B, dark and light blue, respectively), revealed clear effects of the addition of peptide to the sample on the $^1\text{H}-^{15}\text{N}$-HSQC fingerprint (Figure 4.14 B). These strong changes hint at an overall change in the conformational organization of the protein.

![Figure 4.14](image)

Figure 4.14 (A) Overlay of the protein $^1\text{H}-^{15}\text{N}$-HSQC of $^{15}$N-enriched hCaM_S17C after introduction of the dia- and para-magnetic probes consisting of the S17C-linked chelating tag in complex with either the diamagnetic La$^{3+}$ or the paramagnetic Dy$^{3+}$ ion (pink and blue, respectively). (B) Superimposed $^1\text{H}-^{15}\text{N}$-HSQC spectra of the hCaM_S17C linked to the paramagnetic probe in the absence (dark blue) and presence (light blue) of unmodified ERα hinge region peptide.

Altogether these observations indicate the potential of the introduction of the paramagnetic metal ions for the dissection of the detailed intramolecular dynamic rearrangements and the overall orientation of the two lobes with respect to each other in order to investigate the hCaM interaction mode with the ERα peptides as well as other interacting partners.
4.2.4 Protein semi-synthesis towards a hinge region-modified ERα LBD

The influence of hinge region modifications on ERα structure and protein interactions has thus far been studied by employing peptide analogues of the ERα hinge region. The introduction of the hinge region modifications of interest into an extended protein construct, spanning the hinge region and the characteristic α-helix-rich LBD, will provide insights to the potential effects of the PTMs on ligand binding and cofactor interaction. Furthermore, the characterization of hCaM-interaction with such a semi-synthetically modified ERα protein construct will allow to elucidate the potential effects of hinge region-environment on the protein-protein interaction. For this purpose, a protein construct needs to be developed into which the individual PTMs of interest can be homogeneously introduced (Figure 4.15). The method explored here for the controlled and efficient introduction of single as well as different combinations of hinge region modifications is a semi-synthetic approach, in which the differently modified hinge region moieties can be synthesized by Fmoc-SPPS as C-terminal thioester peptides and the C-terminal ERα LBD moiety will be expressed in a recombinant protein construct, yielding an N-terminal cysteine after protein isolation. In this way, the C-terminal ERα LBD can be ligated to differently modified hinge region peptides, yielding a large library of differently labeled proteins (Figure 4.2 D).

![Figure 4.15](image-url) Protein construct of the ERα LBD including the specifically modified N-terminal ERα hinge region.

A crucial detail in the design of the protein constructs is the inclusion of a purification tag at the N-terminus of the fusion protein. In this way, the deliverance of the ERα N-terminal cysteine for ligation would simultaneously remove the purification tag and thereby avoid interference of a purification tag with biophysical characterization of the protein. Furthermore, a C-terminal tag for protein purification would be less favorable due to its potential interference with the C-terminal H12 of the ERα LBD. The envisioned library of modified protein constructs (Figure 4.2) will provide a useful toolbox for the dissection of the influence of individual modifications as well as their cross talk on ligand binding, cofactor interaction, ERα-dimerization and the interplay of these interactions with
the regulatory influences of hCaM. For this purpose a range of different approaches for the introduction of an N-terminal cysteine to the ERα LBD were explored.

**Explored protein constructs**

A similar expression protocol was handled for all described protein constructs. All proteins were expressed in the presence and in the absence of stabilizing agonistic ligand 17-β-estradiol (E2) in TB medium. Target vector transformed *E. coli* BL21 (DE) cells were grown over night in LB medium pre-cultures in the presence of the respective antibiotic substance. Total overnight LB pre-cultures were used to inoculate the 2 L TB expression medium, which was incubated at 37 °C shaking in a baffled 5 L Erlenmeyer flask until a cell density of about 0.8 at A600nm was reached. Bacterial cultures were subsequently cooled to 15 °C, followed by induction of protein expression by IPTG and an overnight shaking incubation for protein expression at 15 °C. Expression cells were separated from growth medium by centrifugation and stored at -80 °C after snap-freezing in liquid nitrogen. Protein degradation was aimed to be avoided by cooling the samples during cell-disruption, and including protease inhibitors in this procedure. Throughout the entire protein purification process, protein solutions, buffers and column materials were cooled to 4 °C or on ice and all procedures were performed in shortest time frames possible to ensure the integrity of the target protein.

Based on the successful introduction of C-terminal PTMs into ERα- and ERβ- LBD (chapter 2), the expression of the ERα LBD with an N-terminal cysteine (CYS_ERα LBD) for the reaction with the modified N-terminal peptide moieties (**Figure 4.2**) was approached by the expression of ERα LBD as fusion protein with an N-terminal self-cleavable Ssp-intein-domain (**Figure 4.16**).

![Figure 4.16](image-url) Schematic overview of the intein-ERα LBD fusion protein construct with an N-terminal chitin binding domain (CBD) for affinity purification.

The protein expression of the intein-ERα LBD construct as described above and subsequent purification via the N-terminal CBD resulted in no detectable amount of neither the fusion protein nor the cleaved ERα LBD. The CBD-intein fusion protein, however was
detected after intein cleavage and subsequent protein elution from the chitin beads affinity resin by SDS-PAGE as well as LCMS analysis (not shown). Since the co-elution of the CBD-intein fusion protein from the chitin bead affinity resin at 4 °C is a common observation and the intein cleavage of the N-terminal Ssp-intein occurs at pH 7, which closely resembles the physiological intracellular conditions during recombinant protein expression, this outcome strongly implies the intracellular auto-cleavage of the Intein-ERα LBD fusion protein during protein expression.

As a result of the observed limitation of the intein-technology, the mechanism of N-terminal cysteine exposure was adjusted. A protease cleavage site was introduced immediately N-terminal of the CYS_ERα LBD in order to yield the CYS_ERα LBD after cleavage of the N-terminal histidine tag (Figure 4.17). Factor Xa protease, originating from the blood coagulating cascade, is known to cleave C-terminal of its recognition sequence with a non-selective promiscuity for the amino acid residue immediately downstream of its recognition sequence (hereafter abbreviated as Xa) and therefore meets the requirements of the protease to yield the CYS_ERα LBD after purification tag cleavage.

![Figure 4.17](image)

**Figure 4.17** Schematic overview of the histidine tagged (His_6) Xa-CYS_ERα LBD expression construct.

Protein expression of the His-Xa-CYS_ERα LBD fusion protein via Ni-NTA affinity chromatography was successful and resulted in reasonably pure protein for subsequent protease cleavage (Figure 4.18 A lane 10 and B lane 1). Furthermore, SDS-PAGE analysis of the protein expression and purification procedure revealed that overnight protein expression induced by IPTG does not yield significant overexpression of the target protein with respect to the abundance observed for the total cell extract before protein induction by IPTG (Figure 4.18 A, lane 4 versus lane 3). It is not evident whether the protein partially resides in the insoluble fraction after cell-opening (Figure 4.18 A, lane 6), which has been observed before for the recombinant expression of ERα LBD constructs. The histagged ERα LBD construct appears to bind well to the Ni-NTA column material (Figure 4.18 A, lane 7 versus 8) and elutes with reasonable purity at an imidazole concentration of 250 mM (Figure 4.18 A, lane 10).
The yield after affinity purification of His-Xa-CYS_ERα LBD remained low at about 1 mg per liter expression culture. Subsequent factor Xa cleavage of the purified fusion protein (Figure 4.18 B) finally yielded a 35 % cleaved protein product (Figure 4.18 B, arrow), as quantified by gel image analysis and confirmed by LCMS (observed masses: 30333 Da for uncleaved and 28120 Da for the cleaved product). This maximum yield was obtained under optimized cleavage conditions at a Xa protease to fusion protein ratio of 1U / 10 μg at 4 °C with a reaction time of 20 hours. Due to the low yield of the protein purification and cleavage, further purification after protease reaction did not yield sufficient amounts of protein for the native chemical ligation reaction with the N-terminal hinge region peptides.

In order to yield higher amounts of target protein, the Xa protease cleavage site was replaced by a TEV protease cleavage site to test whether this would improve the cleavage reaction. Furthermore, an N-terminal GST tag was introduced for affinity purification at milder conditions in view of protein stability as well as potential interference of residual eluant imidazole with the protease reaction. In addition, the recombinant expression of CYS_ERα LBD as GST fusion protein may improve the final protein yield by improving the
overall fusion protein solubility. These changes result in the three subsequent protein constructs (Figure 4.19).

![Figure 4.19 Schematic overview of the histidine tagged (His6) TEV-CYS_ERα LBD (A), GST-tagged Xa-CYS_ERα LBD expression construct (B), and the GST-tagged TEV-CYS_ERα LBD (C).]

The expression and purification of the His6-TEV-ERα LBD (Figure 4.19 A) resulted in very low yields and was not further explored. Protein expression and two-step purification of GST-Xa_CYS_ERα LBD (Figure 4.19 B) by glutathione sepharose affinity chromatography and subsequent size exclusion chromatography (SEC) resulted in similar yields as observed for the His-Xa-CYS_ERα LBD fusion protein (Figure 4.17). In comparison, however, higher protein purity has been achieved in the case of GST-Xa_CYS ERα LBD (Figure 4.20 B, lane 3). The successful isolation of the GST-Xa_CYS ERα LBD fusion protein was confirmed by LCMS (with the expected mass of 62173 Da and the observed mass of 62176 Da).
**Figure 4.20** SDS PAGE analysis of the GST-Xa_CYS ERα LBD (62173 Da) purification by glutathione affinity (A) and size exclusion (B) chromatography as well as subsequent protease cleavage (B, lane 4 and C). Protein expression monitored before and after induction by IPTG (A lanes 2 and 3, respectively), protein abundance after cell- opening in the insoluble (4) and soluble phase (5), flow though off cell extract application to glutathione sepharose (6) and of the subsequent wash steps (7 and 8), non-concentrated elution fraction at 1 mM glutathione (9) and concentrated elution fractions at 5 and 15 mM glutathione (10 and 11, respectively), for reference a protein molecular weight marker in kD (1 and 12). GST-Xa_CYS ERα LBD sample after glutathione sepharose affinity chromatography (B lane 2) and after size exclusion chromatography (SEC) (3) as well as a protein molecular weight marker in kD for reference (1). Samples of SEC-purified fusion protein cleavage by Xa protease before and after cleavage (B, lanes 3 and 4, respectively) with the cleavage products N-terminal GST-tag (34078 Da) and C-terminal CYS_ERα LBD (28113 Da). Optimization of the cleavage conditions with incubation time over night at 4 °C (C) at different ratios of Xa protease (in [U]) over GST-Xa_CYS ERα LBD fusion protein (in µg) of 1 U/10 µg (1), 1 U/1.7 µg (2), 1 U/0.5 µg (3), 1 U/0.25 µg (4). For samples 1 and 2, the amount of Xa protease was doubled after the first hours of incubation. A protein molecular weight marker is shown for reference protein masses in kDa (C, lane 5).

SDS-page image analysis revealed that factor Xa cleavage went to a maximum of 45 % completion under the tested cleavage conditions (**Figure 4.20** C, lane 4). Protease cleavage product CYS_ERα LBD was confirmed by LCMS (28113 Da; observed mass: 28114 Da). Under standard conditions, protein cleavage by factor Xa protease went to 17 % completion as determined by gel image analysis (**Figure 4.20** B, lane 4). For all other tested cleavage conditions a linear increase of about 7 % per sample was observed, correlated to the increase protease excess over protein (**Figure 4.20** C, lanes 1-4). At all conditions with an excess of protease above 1 U per 1.7 µg fusion protein, a lower molecular weight band (approximately 19 kD) was observed and its intensity appears to be correlated with the
relative amount of protease added to the reaction mixture (Figure 4.20 C, lanes 2-4). This band could very well correlate to an off-target digest of the target protein induced by increased levels of factor Xa protease (Figure 4.20 C, lanes 2-4), which is also supported by the increase in in band intensity of the N-terminal GST-tag but a less clear behavior of the band intensities of the CYS_ERα LBD (Figure 4.20 C, lanes 2-4). A designated secondary Xa cleavage site however could not be determined by sequence analysis nor by the resulting masses observed in the LCMS spectra. Again, the low overall yield of final CYS_ERα LBD did not result in a sufficient amount of protein for the establishment of the envisioned modified protein library (Figure 4.15).

**Figure 4.21** SDS PAGE analysis of the protein purification steps of ERα fusion construct GST-TEV-CYS_ERα LBD (63038 Da) by glutathione affinity chromatography (A), SEC-purified fusion protein (B) and TEV protease catalyzed fusion protein cleavage (C). Protein molecular weight marker as reference in kD (A, lane 1 and 12), monitoring protein expression before and after induction by IPTG (2 and 3 respectively), insoluble and soluble phase after cell-opening (4 and 5, respectively), flow through of the glutathione sepharose column material during cell extract application (6), washing steps (7 and 8, respectively) and elution fractions at 1 mM, 5 mM and 15 mM glutathione (9-11, respectively). SEC-purified fusion protein GST-TEV-CYS_ERα LBD (B, lane 2) with protein molecular weight marker as reference in kD (1). Different conditions of overnight TEV cleavage at 4 °C (C, lanes 1-2) and 15 °C (3) of fusion protein GST-TEV-CYS_ERα LBD with 1 U of TEV protease over 1 μg fusion protein (lanes 1 and 3) or 1 U protease over 0.5 μg fusion protein (2), uncleaved fusion protein (4), ~ 10 U TEV protease (5) and protein molecular weight marker as reference in kD (6). The boxes highlight bands appearing upon TEV incubation (C, lanes 1-3).

The effectiveness of the replacement of the Xa cleavage site by a TEV-protease cleavage site was furthermore explored (Figure 4.19 C). As commonly observed for ER expression constructs, protein overexpression of this construct is also not detected by SDS-PAGE analysis (Figure 4.21). However, the elution fractions of the glutathione affinity
puration material reveal the significant abundance of target protein (63.0 kD, Figure 4.21 A, lane 10 and 11). Subsequent purification by SEC resulted in highly pure GST-TEV-CYS_ERα LBD fusion protein (63.0 kD, Figure 4.21 B, lane 2) at similar yields as observed for the analogous Xa-construct.

GST-TEV-CYS_ERα LBD fusion protein cleavage by TEV protease resulted in two additional protein bands, corresponding to the expected molecular weight of the N-terminal GST-tag (34.9 kD) and the C-terminal CYS_ERα LBD (28.1 kD, Figure 4.21 C, lanes 1-3 upper and lower box, respectively). Gel image analysis reveals about 30 % cleavage efficiency and no significant differences between different reaction conditions (Figure 4.21 C, lanes 1-3). Consequently, the replacement of the Xa cleavage site by a TEV cleavage site did not improve cleavage efficiency.

Finally, in order to further on avoid the protease cleavage site for the deliverance of the N-terminal cysteine, a cleavable C-terminal purification tag was introduced to the intein-based protein construct described above (Figure 4.16), resulting in a CBD-intein-CYS_ERα LBD-TEV/Xa-His6 fusion protein (Figure 4.22). Next to representing a compromise on the initially avoided C-terminal purification tag, the cleavable His6-tag provides the possibility for protein detection by antibodies, which could be advantageous for the dissection of problems during protein expression by western blotting as well as the detection during binding assays such as an on-chip cofactor binding array for screening the influence of hinge region modifications on cofactor interaction. As a response to previously observed low yields of target protein expression, an optimization of the protein coding DNA sequence of the CBD-intein-CYS_ERα LBD-TEV/Xa-His6 construct for more efficient target protein overexpression in *E. coli* host cells was performed.

![Figure 4.22 Schematic overview of the CBD-intein-CYS_ERα LBD-TEV/Xa-His6 fusion protein expression construct.](image)

Overexpression and the purification procedure of the CBD-intein-CYS_ERα LBD-TEV/Xa-His6 target protein resulted in no detectable CBD-intein-CYS_ERα LBD-TEV/Xa-His6 fusion protein (55.6 Da), auto-cleaved CYS_ERα LBD-TEV/Xa-His6 (30.4 Da), or the respective CDB-intein fusion protein (25.2 kD) in the elution fractions of the chitin beads or
Ni-NTA affinity chromatography (not shown). Yet, the analysis of the cell-extraction and protein purification procedure of the CBD-intein-CYS_ERα LBD-TEV/Xa-His6 fusion protein by SDS-PAGE (Figure 4.23) revealed a significant abundance of three bands in the insoluble fraction after cell-opening (Figure 4.23, lanes 3a and 3b). These three bands most probably correspond to the CBD-intein-CYS_ERα LBD-TEV/Xa-His6 fusion protein (55.6 kD), the auto-cleaved CYS_ERα LBD-TEV/Xa-His6 (30.4 kD) and the respective CDB-intein fusion protein (25.2 kD). This observation hints at an improved overexpression of the target protein due to codon optimization, since the earlier described fusion proteins have never clearly been distinguishable in whole cell-extract or cell-debris samples. Furthermore a reasonable amount of auto-cleavage prior to fusion protein purification is observed, confirming the observation of intracellular auto-cleavage in earlier expression constructs (Figure 4.16).

![Figure 4.23](image)

**Figure 4.23** SDS-PAGE of the expression at 15 °C (a) and at 22 °C (b) and purification steps of the CBD-intein-CYS_ERα LBD-TEV/Xa-His6 (55561 Da) fusion protein. Bio-Rad double color prestained protein marker in kD (1), supernatant of cell-extract (2), pellet after cell-opening (3), elution fraction 1 and 2 from chitin beads affinity resin (4 and 5, respectively), flow through of cell extract and wash buffer from Ni-NTA affinity resin (6 and 7, respectively).

The observed shift of the entire target protein fraction into inclusion bodies (Figure 4.23), harbors the chance for high-yield and high-purity purification of the target protein under denaturing conditions. However, in that case a reliable refolding protocol must be established in order to retain reasonable amounts of functional protein.
4.3 Conclusion

The described studies of the investigated ERα hinge region peptides revealed that PTMs of this regulatory domain affect regional structural propensity and directly regulate protein-protein interaction by changing the electrostatic properties of the individual amino acid side chains. CD spectroscopy revealed direct effects of lysine acetylations and serine phosphorylation on the stability of the hinge region peptide helical character. Overall, the individual but also combinations of lysine acetylations revealed helix-stabilizing tendencies and the serine phosphorylation significant helix-destabilizing character. The most pronounced stabilizing effect of an individual lysine acetylation was observed for K303. This observation implies that K303-acetylation and S305-phosphorylation make part of two opposing regulatory mechanisms. In fact, this hypothesis is supported by the reported inhibition of K303-acetylation by S305-phosphorylation. Acetylation of K303 has been reported to reduce ERα activation by agonists, which is emphasized by the frequently occurring K303R mutant with increased E2-sensitivity and resulting proliferation of in breast cancer patients. Hyper-phosphorylation of S305 on the other hand has been described as one crucial cause of tamoxifen resistance and enhanced ERα transactivation in breast cancers. The mechanism by which these distal hinge region modifications affect the transactivation of ER has been reported to be linked to altered receptor conformation upon S305 phosphorylation. Furthermore, the hinge region has been reported to play a significant role in mediation of ERα transactivation by regulation of conformational re-orientations of the receptors N-and C-terminal activation functions. The molecular mechanisms underlying these regulatory effects have not been reported to date. The direct effects of hinge region modifications on its structural character as observed by CD spectroscopy provide evidence for a mediation of PTM-effects via direct (de-) stabilizing effects on hinge region helical character. The molecular model, which can be drawn based on these results is that the reduction of helical character by S305-phosphorylation supports a less rigid hinge region and consequently increased rotational freedom of the DBD and LBD, allowing for the concerted action of AF-1 and AF-2, whereas lysine acetylation results in a rigidification of the hinge region and thus an inhibitory effect on AF-1 and AF-2 concertation and receptor transactivation. This initial molecular model could be tested in more detail in the future by a semi-synthetic ER construct spanning the DBD, hinge region and LBD. Even small tendencies as observed here may then be translated to large effects on DNA-binding and
coactivator interaction, mediated by a hinge region lever between the DBD and LBD with a PTM-tuned rigidity.

Next to the intramolecular regulation of ER-transactivation, the investigated PTMs are identified as regulators of protein-protein interaction. The acetylation of K302 and K299, which have been shown to be engaged in direct ionic interactions with hCaM, significantly reduce the affinity towards hCaM, presumably by abolishing the ionic interactions and less so by their effect on hinge region helical character. This observation is in line with the proposed deactivation function of acetylation on ERα activity. In this case, the reduced interaction with hCaM, provides an intramolecular mechanism of ERα-regulation in contrast to the earlier discussed intermolecular mediation. hCaM has been reported to promote ERα activity by scaffolding the formation of ERα-dimers. This stimulation can thus be reduced by the direct regulation of ERα-hCaM-interaction via the described influence of hinge region lysine acetylations. The molecular base of a PTM-regulated hCaM-ERα hinge region interaction should further be characterized by 15N-NMR analysis of the successfully 15N-labeled hCaM and the hinge region peptide analogues. For this purpose, the successful 15N-enrichment of two hCaM cysteine-mutants and introduction of the dia- and para-magnetic probes was described here.

Furthermore, the interaction between a more extended protein construct, spanning the modified ERα hinge region and the LBD is of interest. Such a protein construct also represents a first step towards a semi-synthetic DBD-hinge region-LBD construct for the investigation of hinge region effects in a protein context. For this purpose, a library of fusion protein constructs towards a CYS_ERα LBD native chemical ligation precursor were explored. Taken together, the observations made during the fusion protein development reveal that the introduction of a GST-tag to yield a larger soluble protein fraction was not successful. Factor Xa as well as TEV protease cleavage of the N-terminal purification tags to yield an N-terminal cysteine for NCL did not yield sufficiently high amounts of protease cleavage product, hinting at a sterically hindered protease cleavage site. This problem may be resolved by shifting the ligation site within the protein sequence. The overall yields of purified protein were increased by the addition of agonistic ligand to the expression cultures. However, in the presence of E2, the yield of the individual protein purifications still resided around 1 mg of fusion protein per liter expression medium, which is not sufficient, given the expected loss of protein during the following protease treatment and subsequent final purification step for the removal of protease. In addition, the following NCL reaction requires
another purification step which will again diminish the amount of final protein product. Altogether, a higher initial expression yield of the protein is a necessity for the continuation of the protein semi-synthesis approach. The results of the expression based on an optimized codon sequence may hint at an improvement in the expression yield. A purification from inclusion bodies via the C-terminal cleavable his-tag and the establishment of a protein refolding protocol may provide access to the versatile CYS_ERα LBD NCL precursor. For difficult to express proteins, auto-induction high density shaking cultures has proven successful in different cases and expression as SUMO-fusion protein, is known to increase the solubility of recombinant proteins, which may provide alternative approaches to eventually produce the NCL protein precursor.\textsuperscript{94,95}
4.4 Experimental section

If not mentioned otherwise, all chemicals and materials were used as purchased from the usual sources and without further purification.

**Peptide synthesis.** All peptides were synthesized on Rink amide resin by standard Fmoc-solid phase peptide synthesis (SPPS) either manually or by automated peptide synthesis or a combination of both for the efficient introduction of modified residues. NMP (N-Methyl-2-pyrrolidone) was used as solvent for the entire synthesis protocol. Deprotection of the N-terminal Fmoc-protection group was done in 20 % (v/v) piperidine in NMP. Fmoc-amino acid building blocks were diluted from 200 mM stock solutions to yield a final twofold molar excess of building block over the resin-substitution, reported by the supplier. Coupling reagent HCTU was added at equimolar concentrations and together with 4 equivalents of N,N-Diisopropylethylamine DIPEA. The coupling of each amino acid building block was incubated for 10 minutes. This coupling step was repeated a second time under the same conditions. The remaining free amines after each double coupling step as well as the final the N-terminal amine of the final amino acid of the peptide sequence were capped by acetylation with pyridine : NMP : acetic anhydride (1:3:1: ratio). Peptides were cleaved from the resin by a ratio of TFA : H2O : TIS : EDT (TFA, trifluoroacetic acid; TIS, trisopropylsilane; EDT, 1,2-ethanedithiol) and precipitated in cold diethyl ether during incubation at -20 °C for minimal 1 hour or until precipitation was observed. Successful synthesis of individual peptides was confirmed by analytical LCMS. All peptides were purified by preparative reversed phase HPLC with mass detection in adequate gradients from water to acetonitrile in the presence of 0.1 % TFA, yielding between 1.5 and 6.5 mg of peptide with a purity above 95 % and a purity of about 90 % for peptides (1) and (3). Peptides were stored freeze-dried at -20 °C. For the introduction of the acetyl-lysine residue and the phospho-serine residue, a Fmoc-Lys(Ac)-OH and Fmoc-Ser(PO3BzlH)-OH building block was used, respectively.

**CD spectroscopy.** For the recording of the DC spectra a Jasco J-18 CD spectrometer equipped with a peltier-type temperature controller (PTC-432S) was employed. Spectra were recorded at a peptide concentration of 50 µM in a 2.5 mM sodium phosphate buffer (pH 7.4) either in the absence of 2,2,2-trifluoroethanol (TFE) or in the presence of 10, 30 or 50% TFE in 1 mm path length quartz cuvettes. Quartz cuvettes were washed with 1 % Hellmanex®III solution, H2O, buffer and peptide solution in buffer before measuring. Blanc spectra were recorded for the buffer and used as a baseline for all measurements. Spectra were recorded from 250 to 185 nm with 0.5 nm data pitch, continuous scanning mode, 20 nm/min scanning speed, 2 seconds response time, 0.5 nm band width and accumulation 3. Resulting CD signals were smoothened with Savinsky-Golay, 20 points window, polynomial order 3 and no boundary conditions. Mean residue ellipticity (MRE) was calculated with equation (1):

\[ \text{MRE} = \frac{(M \cdot \theta_\lambda)}{(10 \cdot (N-1) \cdot c \cdot l)} \]  

with the mean residue molar ellipticity (MRE), molecular mass in Da (M), ellipticity in degrees (\(\theta_\lambda\)), the number of amino acids (N), the peptide concentration in g/ml (c) and the path length in cm (l).

**Expression vectors for CaM.** The DNA construct for the expression of human hCaM (hCaM) was designed based on the example of the established hCaM vector (DNA synthesized and sequence optimized for expression in *E. coli* by GenScript). The protein coding sequence was inserted into the expression vector (pET-14b) with the restriction sites NcoI and XhoI yielding the expression vector pET14-b_hCaM. The protein
coding sequence, yielded a tag-free expression of amino acids 1-149 (protein ID: P62158). S17C and A128C mutations were introduced inspired by Anthis et al. for the introduction of metal chelating tags yielding the expression vectors pET14-b_hCaM_S17C and pET14-b_hCaM_A128C (mutagenesis done by GenScript).

Protein expression of hCaM. All flasks, growth media and solutions used for bacterial expression cultures were either autoclaved for 20 minutes at 121 °C or sterilized by filtration. The expression and purification of hCaM was done as described in literature with some adjustments of the protocol. The expression vector pET14-b_hCaM was transformed into E.coli BL21(DE) by classical 30 s heat-shock transformation at 42 °C followed by 2 minute incubation on ice. Subsequently cells were allowed to recover for about two cycles of cell division in Luria-Bertani (LB) medium at 37 °C in the absence of ampicillin. Transformed cells were grown over night on LB-agar plates in the presence of 100 µg/ml ampicillin at 37 °C. Pre-cultures of 50 ml LB medium supplemented with ampicillin 100 µg/ml were inoculated with single colonies of freshly transformed cells and incubated shaking overnight at 37 °C and 220 rpm. 1 L LB medium expression cultures freshly supplemented with 100 µg/ml ampicillin in 2 L baffled Erlenmeyer flasks were inoculated with the entire volume of the overnight pre-cultures and subsequently grown until a cell density at A600nm of 0.8 at 37 °C and 220 rpm. Protein expression was subsequently induced by the addition of isopropyl β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM. Cells were kept on 37 °C and 220 rpm shaking incubation for 4 hours to allow protein expression. The duration of expression time can be doubled for higher protein yield. Protein-expression was finalized by separation of the expression cells from their growth medium by centrifugation for 10 min at 7000 rpm and 4°C. Harvested cells pellets were snap-frozen in N2 (l) and store at - 80 °C until further usage.

Expression of 15N-enriched hCaM mutants. Protein expression for 15N-labeling was performed in 1 L M9 minimal medium (48 mM Na2HPO4·1 H2O, 22 mM KH2PO4, 8.6 mM NaCl, 18.6 mM NH4Cl, 2 mM MgSO4, 0.4% glucose, 2 mM CaCl2) supplemented with 100µg/ml ampicillin. During autoclaving of the M9 medium, salt precipitation was observed. Consequently the medium was not autoclaved prior to protein expression but filtered sterile. The expression vectors pET14-b_hCaM_S17C and pET14-b_hCaM_A128C were transformed into E. coli BL21(DE) cells individually. Single colonies from LB agar plates, supplemented with 100 µg/ml ampicillin were transferred to individual 25 ml LB medium with 100 µg/ml overnight pre-cultures at 37 °C and shaking at 220 rpm. Entire pre-cultures were used to inoculated 1 L rich TB expression medium (24 g/L yeast extract, 12 g/L peptone, 4 ml/L glycerol) supplemented with separately autoclaved TB buffer (9.4 g/L K2HPO4, 2.2g/L KH2PO4) in the presence of 100µg/ml ampicillin. Cells were grown at 37 °C and shaking at 220 rpm up until an optical density of A600nm ~ 1 and were subsequently separated from the medium by centrifugation at 6000 rpm for 10 minutes at 4 °C. Cells were transferred to 1 L M9 minimal medium and re-suspended in the culture shaking incubator at 37 °C and 220 rpm. Subsequently, IPTG was added to a final concentration of 0.5 mM and cells were incubated at 37 °C and 220 rpm over night for protein expression. Cells were harvested by centrifugation at 7500 rpm and 4 °C, snap-frozen in N2(l) and stored at -80°C until protein purification.

Protein purification of the different hCaM constructs. The entire purification procedure of the different hCaM constructs was performed at 4 °C and cooled buffer solutions whenever possible. Previously harvested expression cells were thawed from -80 °C on ice. Cell lysis was performed by ice water-cooled homogenization (Emulsiflex C3, Avestin) at a minimum of two cycles in lysis-buffer with 50 mM Tris-HCl at
pH 7.5, 2 mM EDTA, 5 mM MgCl₂ and freshly added 200 μM PMSF and 1 μg/ml DNAse. During the purification procedure of the cysteine mutants all buffers and protein solutions were supplemented with 1 mM TCEP. Subsequent centrifugation for the separation of cell debris at 20000 rpm and 4 °C for 30 minutes yields the cell extract with the soluble target protein in the supernatant. Prior to gravity column chromatography protein purification, the protein was saturated with calcium by adding CaCl₂ to the cell extract at a final concentration of 5 mM. Column chromatography was performed with Phenyl-functionalized sepharose (Phenyl-sepharose CL4B, Sigma) with a column volume (CV) of 15 ml per 1 L expression culture. Column material was equilibrated with 5 CVs of 50 mM Tris-HCl at pH 7.5 with 5 mM CaCl₂ and 0.1 M NaCl. Cell extract was flown over the pre-equilibrated column material by gravity flow. Unbound protein was washed off the column with a buffered solution of 50 mM Tris-HCl at pH 7.5 with 0.1 mM CaCl₂ and 0.1 M NaCl until the flow through reached baseline absorption. A second wash step at a higher salt concentration was performed (50 mM Tris-HCl at pH 7.5 with 0.1 mM CaCl₂ and 0.5 M NaCl) until baseline absorption was reached. Protein was eluted from the column material by addition of calcium-chelating agent EGTA (50 mM Tris-HCl at pH 7.5 with 1 mM EGTA). The eluted protein was concentrated and buffer was exchanged to the storage buffer (20 mM Tris-HCl at pH 7.5 with 2 mM CaCl₂ and 150 mM NaCl) by multiple cycles of concentration in Amicon® Ultra 3K centrifugal filter devices. Remaining theoretical concentrations of EGTA were below 1 μM. For higher purity of the protein and additional size exclusion chromatography step (HiLoad 26/60 Superdex 75 column, GE Healthcare) in the storage buffer conditions was performed.

**Introduction of dia-/paramagnetic metals and their chelating tag to the hCaM cysteine mutants.**

The design of the metal chelator tag was inspired by previously reported EDTA-based chelators. The different metals were diluted from 2 mM stock solutions to yield a 1:1 ratio of metal : chelator tag (Figure 4.24) in H₂O. These two components were pre-incubated and monitored by NMR. The saturated metal-chelator complex was subsequently added to the hCaM-mutants in hCaM-storage buffer (see section ‘Protein purification of the different hCaM constructs’). The reaction was monitored by NMR at a 1 : 1.1 ratio of the metal-chelator complex to protein.

**Solution NMR of hCaM.** NMR samples of the different ¹⁵N hCaM proteins were prepared in TRIS buffer (20mM TRIS, 150mM NaCl, 5mM CaCl₂, pH 7.5) to a final concentration of 150μM, containing 10% D₂O. ¹H¹⁵N HSQC (heteronuclear single quantum correlation) spectra were collected at 37°C on a Bruker 900 MHz spectrometer (except for the hCaM_wt, which was recorded at 750 MHz). HSQC spectra were acquired with 1024 complex data points in the ¹H dimension (t2) and 180 or 256 increments in the indirect dimension (t1). Data sets were zero-filled to 2048x512 before Fourier transformation using Topspin (Bruker), and analyzed
with CARA 1.8.4. The chemical shifts of hCaM in the bound state, as assigned by Zhang et al.\(^\text{40}\) were used as reference for the backbone assignment.

**Expression vectors for CYS\(_{ERα}\) LBD.** Integrity of all used expression vectors was confirmed by sequencing. The expression vector for the His\(_{56}-\text{Xa-CYS\(_{ERα}\) LBD}\) construct (pQE-30- His\(_{56}-\text{Xa-CYS\(_{ERα}\) LBD}\) S309C-554) was constructed by standard molecular cloning. ER\(_{α}\) LBD domain coding DNA sequence was amplified from the pTWIN1-CBD-intein-S309C\(_{ERα}\) LBD vector previously cloned by Sabine Moecklinghoff in a classical PCR reaction with the primers 5'-G T C T C A G C A G C T G A C G G C C G A C C A G A T G G-3' and 5'-G T A C A T G A A G C T T G C T A G T G G G C G C A T G T A G G-3'. The choice of the primer sequences allows for the digest with the restriction enzymes PvuII and HindIII and the subsequent introduction of the amplified DNA-insert into the target vector pQE-30 Xa (Qiagen), which was double digested with HindIII and Stul, resulting in a blunt end ligation site for the PvuII blunt end digested DNA-insert. Furthermore the choice of the latter restriction enzymes results in the loss of the codon for the N-terminal cysteine 309, required for the native chemical ligation reaction. Cysteine 309 was subsequently reintroduced via site-directed mutagenesis following the quick change protocol via classical PCR with the partially non-complementary primer pair 5'-C G A G G G A A A T T A T T T T C A G T G C C T G A C G G C C G A C C-3' and reverse primer 5'-C T G A A A A T A C A G G T T T T C C T C C T C C T C G A T A C C A G A T C-3'.\(^{98}\)

The expression vectors for the fusion protein constructs GST-Xa-CYS\(_{ERα}\) LBD, GST-TEV-CYS\(_{ERα}\) LBD were generated by amplification of the CYS\(_{ERα}\) LBD S309C-554 coding DNA-fragment from the pQE-30-His\(_{56}-\text{Xa-CYS\(_{ERα}\) LBD}\) construct by classical PCR with the forward primer 5'-G G A G A A A A C C T G T A T T T T C A G T G C C T G A C G G C C G A C C-3' and reverse primer 5'-C T G A A A A T A C A G G T T T T C C T C C T C C T C G A T A C C A G A T C-3'. The expression vectors for the C-terminal His-tagged CDB-intein-CYS\(_{ERα}\) LBD S309C-554-His construct was ordered at GenScript with an optimized codon sequence for the expression of the target protein in *E. coli* expression cells. The CYS\(_{ERα}\) LBD sequence was in this case cloned into a pET-52b vector with the restriction sites BamHI and SalI, yielding the expression vector pET-52b-CDB-intein-CYS\(_{ERα}\) LBD-His.

**Protein expression of the different CYS\(_{ERα}\) LBD fusion constructs.** All flasks, growth media and solutions used for bacterial expression cultures were either autoclaved for 20 minutes at 121 °C or sterilized by filtration. The protein expression of the different CYS\(_{ERα}\) LBD (S309C-554) protein fusion constructs was generally done in *E. coli* BL21 (DE) cells overnight at 15 °C and shaking incubation at 220 rpm. Variations on the expression protocol such as change in expression temperature and duration or amount of IPTG used for expression induction were implied, yet did not significantly alter expression outcomes. 25 ml LB pre-cultures, supplemented with the respective antibiotic, were inoculated with freshly transformed cell colonies or from
glycerol stocks, stored at -80 °C. Overnight bacterial growth at 37 °C under 220 rpm shaking incubation in baffled Erlenmeyer flasks was followed by a sterile transfer of the entire pre-culture volume into 2 L TB medium expression media with the respective antibiotic. Cells were allowed to grow until an A₆₀₀nm of about 0.8 while incubation at 37 °C shaking at 220 rpm in a baffled 5 L Erlenmeyer flask. Bacterial cultures were subsequently cooled to 15 °C in ice water, followed by induction of protein expression by 0.1 mM IPTG. In case of protein expression in the presence of ligand, TB expression medium was supplemented with 10 µM E2 upon induction with IPTG. Protein expression was allowed to proceed overnight at 15 °C. Expression cells were separated from growth medium by centrifugation for 10 minutes at 7000 rpm and 4 °C and stored at -80 °C after snap-freezing in liquid nitrogen.

**Protein purification of the different CYS_ERα LBD fusion constructs.** Throughout the different protein purification procedures, protein solutions, buffers and column materials were cooled to 4 °C or on ice and all procedures were performed in shortest time frames possible to ensure the integrity of the target protein. Cell-lysis was performed by 2-3 cycles of ice-water cooled homogenization (EmulsiFlex C3, Avestin) in the respective starting buffer of the subsequent purification procedure, supplemented with 200 µM phenylmethanesulfonyl fluoride (PMSF), 5 mM MgCl₂ and 1 µg/ml DNAse. Soluble cell-extract was subsequently separated from cell-debris by centrifugation for 30 minutes at 20000 rpm and 4 °C and applied to pre-equilibrated column chromatography resins. Affinity purification via N-terminal and C-terminal histidine purification tags was done by Ni-NTA column chromatography over either ~15 ml CV of charged gravity flow His rind resin (Novagen) or by FPLC on 10 ml CV pre-packed HisTrap FF (2 x 5 ml, GE Healthcare Life Sciences) per 2 L expression culture. In both cases, column material was pre-equilibrated in 5 CV of 1 x PBS at pH 8 in 370 mM NaCl and 40 mM imidazole, after cell-extract application washed in 5 CV of 1 x PBS at pH 8 in 370 mM NaCl and 40 mM imidazole and eluted by a gradient of increasing imidazole concentration by either a step-wise gradient (100 mM, 250 mM and 500 mM imidazole) or a linear gradient up to 500 mM imidazole in 1 x PBS at pH 8 in 370 mM NaCl. GST-tagged proteins were purified by affinity column chromatography via glutathione sepharose (5 ml GSTrap FF columns, GE Healthcare Life Sciences per 2 L expression culture). Buffers were prepared according to the recommendations of the column supplier and elution of the target protein from the affinity chromatography resin was done by a step-wise increase of reduced glutathione (5, 10 and 15 mM). SEC was performed on a HiLoad 26/60 Superdex 75 column (Ge Healthcare) in 20 mM Tris-HCl at pH 6.5 with 50 mM NaCl, 1 mM CaCl₂.
4.5 Acknowledgements

Christian Haase synthesized a hinge region peptide intended for the ligation to the CYS_ERα LBD NCL precursor for the introduction of S305 phosphorylation. The hinge region peptide CD studies and most of the work towards a CYS_ERα LBD NCL precursor is based on the Master project of Anneloes Oude-Vrielink. Lidia Nieto performed the ITC and NMR measurements. The synthesis and design of the EDTA-based metal-chelator tag was done by Ángeles Canales and co-workers. Sabine Moecklinghoff cloned the pTWIN1-CBD-intein-CYS_ERα LBD-construct during the period of her PhD project. Ralph Bosmans and Chan Vinh Lam helped with the QTOF-access under the supervision of Joost van Dongen.
4.6 References

(1) Kumar, V.; Green, S.; Stack, G.; Berry, M.; Jin, J. R.; Chambon, P. Cell 1987, 51, 941–951.


Estrogen Receptor Hinge Region Modifications and their Influence on the hCaM-Interaction


(66) Nieto, Lidia; Tharun, Inga; Canales, Angeles; Mallagaray, Alvaro; Perez-Castells, Javier; Jimenez-Barbero, Jesus; Brunsveld, Luc. Structural insights to hCaM interaction with estrogen receptor alpha hinge region and its modulation by post-translational modifications, Manuscript in preparation.


5 Molecular insights into the interaction between
14-3-3 proteins and the ERα F-domain

Abstract. Resistance to endocrine therapy is a commonly encountered limitation in breast cancer therapy. The concept of estrogen receptor (ER) inhibition in ER-positive breast cancer, however, remains a promising goal and can be beneficial for large patient groups. Therefore alternative approaches on ER inhibition have evolved, amongst which the targeted inhibition of ER protein-coactivator interaction. In contrast to inhibiting stimulating protein-protein interactions, the concept of stabilizing inhibiting complexes with natural binding partners has recently been developed. Ubiquitous 14-3-3 proteins provide a promising protein scaffold for this purpose. In order to understand the molecular determinants of the recently discovered phosphorylation-dependent interaction between ERα and 14-3-3 proteins, the protein interaction was quantified by biophysical methods, using peptide analogues of the C-terminal ERα F domain. The ERα F domain has not been studied in great detail yet and new insights on domain characteristics are described in this chapter. Furthermore, native chemical ligation was applied for the introduction of the F-domain phosphorylation to an ERα protein construct. This protein construct is aimed to the biophysical characterization of the F-domain and its modification on intra- and inter-molecular modulations and their potential role in the development of new protein-protein interaction stabilizers.
5.1 Introduction

The most C-terminal domain of ERα, the ~40 amino acid residue F-domain, represents one of the pronounced structural differences between the two ER subtypes α and β and has consequently been proposed to play a role in the differential regulation and function of the two ER subtypes. ERα F-domain function and structure are not well-characterized and differ greatly from the F-domain of other nuclear receptors which makes analogies improbable. ERα F-domain is involved in the control of receptor dimerization and plays a role in mediating ligand effects including the agonistic effects of tamoxifen on ERα. The structural properties of the F-domain are not yet understood in detail. The most direct description of F-domain structure so far originates from the co-crystallization of an F-domain peptide analogue in complex with 14-3-3σ, which reveal a largely extended conformation for the five C-terminal F-domain residues. However, structural data of the full F-domain e.g. by crystallization as a C-terminal extension of the prevalently crystallized ERα ligand binding domain (LBD) is not available. The F-domain is consequently considered to be largely unstructured, yet alternative methods like in silico structure predictions have led to the assumption that the F-domain sequence covers β-strand as well as α-helical characteristics. Independently of the feasibility of the predicted structural character, residues embedded in the supposedly structured regions have been shown to affect F-domain regulatory character, such as altering ligand effects, emphasizing the importance of F-domain regulatory function. Taken together it is evident that the ERα F-domain exerts important regulatory function on the receptors activity and must be considered in the development of ER-targeted drugs, due to its involvement in mediation of ligand effects. Furthermore, the role of the F-domain in fine-tuning ERα function has been reported to be modulated by post-translational modifications (PTMs). PTMs of the F-domain, affecting DNA- and ligand-interaction, have been described to occur in a cellular sub-population of ERα in the N-terminal sequence region of the F-domain and thus in close proximity to the functionally important helix 12 (Chapter 2). Recently, the novel phosphorylation site of a conserved penultimate threonine (pT595), has been identified and phosphorylation-induced protein-protein interaction between pT-ERα and 14-3-3 has been characterized. Abrogation of the interaction between 14-3-3 and ERα by mutation of T595 is sufficient to enhance ERα transactivation, yet once stabilizing the protein-protein-interaction with the well-established 14-3-3 interaction stabilizer fusicoccin (FC), E2-induced transcription and cell-growth is reduced significantly. Antagonist 4-hydroxy-tamoxifen as well as agonist E2 have been shown to enhance ERα
interaction with 14-3-3 in a threonine-phosphorylation dependent and ERα subtype specific manner. 9 14-3-3 proteins are highly conserved ubiquitous phospho- serine or threonine recognizing adaptor proteins, which are forming protein-protein complexes with a vast number of proteins and are thereby often involved in the regulation of their activity in vital processes such as metabolism, signal-transduction and cell-cycle regulation. 13 Consequently they have been associated with the pathophysiology of diverse diseases, amongst which neurodegeneration and several cancer types. 14–17 In fact, enhanced expression of 14-3-3ζ is stimulated by tamoxifen and has been associated with poor clinical effect of endocrine treatment of breast cancer patients with the anti-estrogen tamoxifen and an early onset of breast cancer relapse. 18,19 Resistance to endocrine treatment is a commonly occurring problem in treatment against estrogen receptor dependent breast cancer. The approaches to finding alternative drug targets have focused for example on the inhibition of ER transcription factor activity by targeting the ER coactivator interaction or by the selective down-regulation of ERα by inducing its degradation. 20–24 The recently described phosphorylation-induced interaction between ERα and 14-3-3 revealed a new protein-protein interface which can be explored as a potential drug target in endocrine therapy for resistant breast cancer patients. 25 Stabilization of the 14-3-3 –ERα interaction presents a promising strategy for the inhibition of ERα nuclear transcription factor activity by the inactivation of ERα through complexation with 14-3-3. However, since 14-3-3 has a vast number of cellular interaction partners, creating specificity for the 14-3-3-ERα interaction stabilization constitutes one of the major challenges of this approach. In order to design effective and specific interaction stabilizers, the molecular determinants and characteristics of the 14-3-3 ERα protein –protein interaction need to be further elucidated in detail. The introduction of the threonine-phosphorylation of interest into a domain fusion construct, spanning the native sequence of the ERα LBD, followed by the full-length F-domain, by native chemical ligation will provide a useful tool for the investigation of the detailed molecular, structural and functional effects of the F-domain in the first place and simultaneously for the effect of threonine-phosphorylation thereon. Furthermore, the in vitro characterization of the interaction between 14-3-3 subtypes and the ERα LBD-F-domain construct will allow for a biophysical characterization of the protein-protein interaction, which will extend our knowledge and provide information for the rational design of 14-3-3-ERα modulators.
Chapter 5

5.2 Results and discussion

5.2.1 Quantification of the interaction between 14-3-3 and pT-ERα

The protein-protein interaction between threonine-phosphorylated (pT) ERα F-domain peptides and different 14-3-3 subtypes and mutants was characterized in a direct in-solution fluorescence polarization (FP) assay. The biophysical characterization allows the dissection of the effects of F-domain pT on the protein-protein interaction as well as the influence of a well-characterized 14-3-3 protein interaction stabilizer fusicoccin (FC) on this interaction.26–28 Since it is known that the lengths of the peptide binder can affect the 14-3-3 interaction with its target proteins by possibly allowing additional stabilizing interactions,9,29 three different lengths of the F-domain region were N-terminally fluorescently labeled for detection and tested for their binding affinity towards the 14-3-3 subtypes σ and ζ as well as the C-terminal deletion mutants of the two proteins (ΔC) (Table 5.1). The C-terminal deletion mutants are of interest for this interaction characterization because they are often used in crystallography studies.9

Table 5.1 Protein variants and peptides, used in in the direct interaction assay for the 14-3-3 -ERα interaction characterization with the phosphorylated threonine indicated as “pT”.

<table>
<thead>
<tr>
<th>Protein name</th>
<th>Protein sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>14-3-3ζ</td>
<td>MSYYHHHHHHHDYDIPTTENLYFQGAMGS-M1-S248</td>
</tr>
<tr>
<td>14-3-3ζ-ΔC</td>
<td>GAMGS-M1-T231</td>
</tr>
<tr>
<td>14-3-3σ</td>
<td>MSYYHHHHHHHDYDIPTTENLYFQGAMGS-M1-N245</td>
</tr>
<tr>
<td>14-3-3σ-ΔC</td>
<td>GAMGS-M1-S230</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Peptide name</th>
<th>Peptide sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>pT-8mer</td>
<td>FITC-A588EGFPApTV595</td>
</tr>
<tr>
<td>pT-15mer</td>
<td>C(FITC)K581YYITGEAEGFPApTV595</td>
</tr>
<tr>
<td>T-15mer</td>
<td>C(FITC)K581YYITGEAEGFPATV595</td>
</tr>
<tr>
<td>pT-29mer</td>
<td>C(FITC)H567LATAGSTSSHSLLQIKYYITGEAEGFPApTV595</td>
</tr>
</tbody>
</table>
For the FP direct interaction assay, the concentration of fluorescent peptide was kept constant and the different 14-3-3 variants were serially diluted over a range from high micromolar to low nanomolar concentrations (Figure 5.1).

**Figure 5.1** Fluorescence polarization assay with different truncation lengths of the ERα F-domain binding site for 14-3-3 proteins. The truncation sequences include threonine-phosphorylated (pT) peptide FITC - A₅₈₈ E G F P A pT V₅₉₅ (pT-8mer), C (FITC) - K₅₈₃ Y Y I T G E A E G F P A pT V₅₉₅ (pT-15mer) and C (FITC) - H₅₆₇ L A T A G S T S S H S L Q K Y I T G E A E G F P A pT V₅₉₅ (pT-29mer) as well as the non-phosphorylated T-15mer. Binding of all different truncation sequences is tested to the sub-types 14-3-3ζ (A and B) and 14-3-3σ (C and D), also including the C-terminal deletion mutants of both sub-types (ΔC, B and D). All interactions were tested in the absence (empty symbols) and presence (filled symbols) of fusicoccin (FC). Concentration of the fluorescent probe was kept constant at 20 nM, FC was added at a constant concentration of 10 µM, yielding a final DMSO concentration of 0.1 % which was also included in the ligand-free samples.
The direct interaction study by FP reveals overall highly similar results for the individual 14-3-3 variants (Figure 5.1 A-D). The FP assay confirms the expected necessity of threonine-phosphorylation for the effective protein-peptide interaction, as demonstrated by the inclusion of non-phosphorylated T-15mer peptide (Figure 5.1 A-D, gray squares). For all different truncation lengths of F-domain peptides a significant shift of the binding curves towards lower protein concentrations was observed upon the addition of FC (Figure 5.1 A-D, empty versus filled symbols), indicating higher binding affinities in the presence of FC. Furthermore, different FP levels at saturating protein concentrations were observed for the three different truncation lengths with a high FP saturation plateau for the 15mer peptide, an intermediate FP saturation for the 8mer peptide and a low FP saturation plateau for the 29mer peptide (Figure 5.1 A-D, gray circles, black triangles and black circles, respectively). The low FP saturation levels observed for the 29mer peptide may be caused by a relatively long N-terminal sequence stretch, protruding from the 14-3-3 binding site and resulting in a high degree of rotational freedom of the fluorophore and thus low FP saturation levels. This observation could indicate that no additional N-terminal interactions are established in this specific elongated peptide sequence, which would otherwise probably result in a more fixed orientation and thus higher FP saturating levels. Potential additional interactions with even longer peptide analogues remain to be investigated. Interestingly, the FP saturation levels could not be directly correlated to the length of the pT-peptides, since overall higher FP saturation levels were observed for the 15mer than for the 8mer peptide. The molecular origin of this observation remains unclear. The saturation FP levels of the individual peptides were not largely affected by the addition of FC, which is in line with the involvement of the C-terminus of the peptides in the direct physical stabilization by FC. This observation hints at a very local effect of FC and at an absence of direct or indirect effect on the dynamics of the peptide N-terminus. The strength of the stabilizing effect of FC was in addition demonstrated by the induction of low, yet significant association between the T-15mer peptide and all 14-3-3 variants at high 14-3-3 concentrations (Figure 5.1 A-D, filled versus empty gray squares), which is in line with the observed residual 14-3-3 effect on non-phospho mutants in a cellular context. The quantification of the obtained FP data yields the $K_D$ values listed in Table 5.2.
Table 5.2 K_D values in nM obtained from FP binding assays (Figure 5.1), quantify the interaction between 14-3-3 variants and different ERα-F-domain truncation sequences (Table 5.1).

<table>
<thead>
<tr>
<th></th>
<th>pT-8mer</th>
<th>T-15mer</th>
<th>pT-15mer</th>
<th>pT-29mer</th>
</tr>
</thead>
<tbody>
<tr>
<td>(-) FC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14-3-3ζ</td>
<td>821 ± 52</td>
<td>- a</td>
<td>97 ± 4</td>
<td>121 ± 11</td>
</tr>
<tr>
<td>14-3-3ζ-ΔC</td>
<td>189 ± 12</td>
<td>- a</td>
<td>55 ± 2</td>
<td>65 ± 5</td>
</tr>
<tr>
<td>14-3-3σ</td>
<td>1310 ± 85</td>
<td>- a</td>
<td>220 ± 8</td>
<td>358 ± 34</td>
</tr>
<tr>
<td>14-3-3σ-ΔC</td>
<td>661 ± 40</td>
<td>- a</td>
<td>226 ± 15</td>
<td>150 ± 11</td>
</tr>
<tr>
<td>(+) FC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14-3-3ζ</td>
<td>17 ± 1</td>
<td>- a</td>
<td>2.4 ± 0.1</td>
<td>25 ± 1</td>
</tr>
<tr>
<td>14-3-3ζ-ΔC</td>
<td>12 ± 1</td>
<td>- a</td>
<td>2.4 ± 0.2</td>
<td>27 ± 1</td>
</tr>
<tr>
<td>14-3-3σ</td>
<td>61 ± 3</td>
<td>- a</td>
<td>5.4 ± 0.1</td>
<td>33 ± 1</td>
</tr>
<tr>
<td>14-3-3σ-ΔC</td>
<td>80 ± 7</td>
<td>- a</td>
<td>7.9 ± 0.3</td>
<td>34 ± 2</td>
</tr>
</tbody>
</table>

a no binding (saturation) in tested concentration range

The quantitative analysis of the FP assay, confirmed the overall increase in binding affinities in the presence of FC (Table 5.2, top versus bottom), emphasizing the stabilizing effect of FC on the protein-peptide interaction. The effect of FC is similar for the pT-15mer and pT-8mer peptides, yet significantly smaller for the pT-29mer peptide (Table 5.2). In the case of the pT-8mer and pT-15mer peptides, a significantly larger stabilizing effect was observed for the wild type 14-3-3 variants when compared to their respective C-terminal deletion mutants. Taken together, the overall binding affinities of the pT-15mer peptide in the presence of FC are significantly higher than the binding affinities of the other two truncation sequences, which is in line with reported data. The contribution of peptide lengths to the (de-) stabilization of the protein-peptide interaction was observed by a significant increase in binding affinity for the pT-15mer peptide with respect to the pT-8mer peptide, hinting at additional stabilizing interactions of the longer pT-15mer peptide. The main difference between the pT-8mer and the pT-15mer is the inclusion of an N-terminal sequence stretch, which is highly conserved in the ERα amongst different species (K581YYIT, Table 5.1) and may hint at the importance of this conserved region. In contrast to the stabilizing effect of the
elongation to 15 amino acids, the elongation of the ERα F-domain analogue to a pT-29mer peptide appeared to have little effect on the binding affinities in the absence of FC and an overall destabilizing effect in the presence of FC (Table 5.2). A stabilizing interaction between the His577 (Table 5.1), which is included in the pT-29mer, in analogy to the 14-3-3-plant ATPase interaction cannot be concluded from this data. Co-crystallization with an elongated pT-ERα peptide or the pT-ERα LBD-F-domain (vide infra) will provide clarity on the detailed protein-peptide or protein-protein interface. A crystallization approach with an elongated F-domain peptide could possibly elucidate the potential back-folding of the N-terminally elongated sequence for additional interactions with the 14-3-3. A co-crystallization between pT-ERα LBD-F-domain would yield even more holistic information about the potential impact of the LBD-fold on the F-domain structural orientation and may furthermore reveal the stoichiometry of 14-3-3-ERα complex composition.

Another correlation that has been observed during the FP interaction characterization (Table 5.2) is the increase of binding affinity towards the C-terminal deletion mutants 14-3-3ζ-ΔC and 14-3-3α-ΔC with respect to the full-length 14-3-3 proteins in the absence of FC (Table 5.2, top). In the presence of FC, however, no such overall tendency has been observed. In fact in the presence of FC, the C-terminal deletion has much smaller effects on the peptide binding affinities (Table 5.2, bottom). In most conditions that were tested, C-terminal deletion has a slightly destabilizing effect on the protein-peptide interaction in the presence of FC. The nanomolar binding affinity, determined for the interaction between 14-3-3ζ with the pT-15mer peptide (Table 5.1) in the absence of FC, was confirmed by surface plasmon resonance experiments (K_D = 265 ± 20 nM, data not shown), confirming the relevance of the reported data.

5.2.2 Protein semi-synthesis of pT-ERα LBD

Introduction of the penultimate F-domain phospho-threonine to an ERα protein construct, spanning the receptors ligand binding domain (LBD) as well as its -F-domain (F), by protein semi-synthesis will provide the possibility to investigate potential stabilizing or destabilizing effects of ERα LBD-F-pT binding to the different 14-3-3 subtypes in molecular and atomic detail. For the design of the semi-synthesis approach by native chemical ligation (NCL), native cysteine residues within the protein sequence are preferably used as ligation junctions. The most C-terminal cysteine of the ERα protein sequence, however, is located 65 amino acid residues upstream of the protein C-terminus. In order to achieve high yield
peptide synthesis, the lengths of the synthetic moiety was reduced to 30 amino acids. The latter was done by introducing the mutation S566C into the final protein sequence to create a de novo ligation junction (Figure 5.2). This approach was inspired by previous studies on which F-domain length could stably be expressed as a C-terminal extension of the ERα LBD in bacterial recombinant protein expression systems, since the recombinantly expressed ERα LBD-F- fusion construct has been described to be subject of degradation.30

![Figure 5.2](image.png)

**Figure 5.2** Protein semi-synthesis of a phosphorylated ERα-LBD-F-domain fusion construct by native chemical ligation (NCL) of the recombinantly expressed N-terminal ERα LBDK302-Q565-thioester (left) and the threonine-phosphorylated (pT) F-domain peptide (right), synthesized by Fmoc solid phase peptide synthesis (SPPS).

A protocol, previously established for the expression and purification of the ERα LBDK302-Q565-Mxe intein-chitin binding domain (CBD) fusion protein was followed in a first approach towards thioester purification for the NCL approach (Figure 5.2).30 Protein overexpression overnight at 15 °C and subsequent purification by chitin bead affinity column chromatography with an on-column intein cleavage step yielded 40 mg reasonably pure protein thioester per liter expression medium (Figure 5.3 A, lane 7).
Figure 5.3 SDS-PAGE analysis of ERα LBD<sub>K302-Q565</sub>- Mxe intein- CBD fusion protein (61.6 kD) purification and intein cleavage yielding the ERα LBD<sub>K302-Q565</sub> and CBD-intein moieties (33.8 kD and 27.9 kD, respectively). Showing a sample of cell extract (1), flow through of cell extract application to chitin beads affinity resin (2), washing off unbound protein (3), buffer exchange of column material to cleavage conditions (4), flow through (5) and supernatant (6) of molecular weight cut-off membrane, final protein solution (7) and protein molecular weight marker (8).

SDS-PAGE analysis of the ERα LBD<sub>K302-Q565</sub>-thioester purification (Figure 5.3) confirms that the fusion protein is immobilized on the affinity column and that the transition to cleavage conditions immediately induces intein-cleavage as indicated by two protein bands around 30 kD (Figure 5.3, lane 4). Co-elution of ERα LBD<sub>K302-Q565</sub>- Mxe intein- CBD fusion protein was also observed under these cleavage buffer conditions. Concentration of the elution fractions after overnight intein cleavage (lane 6) confirms the specific enrichment of the cleaved ERα LBD<sub>K302-Q565</sub> (33.8 kD). The protein band corresponding to a smaller molecular weight, which was observed in the product fraction (lane 6 and 7) can be assigned to the CBD-intein fusion domain (27.9 kD) and the protein band of higher intensity around 60 kD can be assigned to residual uncleaved ERα LBD<sub>K302-Q565</sub>- intein-CBD fusion protein (61.6 kD). The vague band in the flow through of the molecular weight cut off concentrator (lane 5) most probably originates from contamination with sample 6. Reasonable purity after one-step affinity purification of about 80 % was quantified by gel image analysis (lane 7).<sup>31</sup> In order to remove the residual CBD-intein fusion protein, observed in the purified thioester (Figure 5.3, lane 7), the protein thioester was incubated with fresh chitin beads prior to the NCL reaction. The protein thioester was subsequently engaged in a small-scale ligation reaction with the F-domain peptide S<sub>S66</sub>CHLATAGSTSSHQKYITGEAEFGPApTV<sub>S95</sub>. To ensure the reduction of the sulfhydryl group of the N-terminal cysteine,
the F-domain peptide was pre-incubated at room temperature with 5 mM tris (2-carboxyethyl) phosphine hydrochloride (TCEP) in ligation buffer (20 mM Tris-HCl at pH 8.0 with 500 mM NaCl and 1 mM EDTA). A threefold molar excess of peptide over protein was included in the ligation reaction at 4 °C overnight in ligation buffer supplemented with 130 mM sodium 2-mercaptoethanesulfonate (MESNA) and 1.5 mM TCEP. The successful ligation, yielding the threonine-phosphorylated ERα LBD-F-domain fusion product (ERα LBD-F-pT) was confirmed by LCMS chromatography (Figure 5.4).

A peak, corresponding to the CBD-intein fusion protein was not observed after ligation (Figure 5.4), indicating that additional incubation with chitin beads prior to the ligation reaction successfully removed the co-eluted fusion protein. The small scale ligation product was separated from the residual unreacted F-domain peptide by means of size exclusion and included in co-crystallography screens with 14-3-3σ-ΔC. Unfortunately no crystal growth was observed. During the up-scaling of the ligation reaction in order to extend the crystallization condition screening and with the intention to characterize the protein-protein interaction in more detail, thioester hydrolysis hindered the formation of ligation product. Repeated trials to reproduce the successful thioester purification failed. Hydrolysis of the thioester was observed in all following attempts already within the time frame of the overnight on-column intein cleavage step and was also observed when the intein cleavage reaction time was reduced. Various alterations in the purification and cleavage conditions, such as the decrease in pH or the increase in MESNA concentration in order to stabilize the thioester, repeatedly resulted in the hydrolysis of the protein thioester, indicating that the ERα
LBD_{K302-Q565}-thioester is very labile. In order to avoid on column hydrolysis of the ERα LBD_{K302-Q565}-thioester, the ERα LBD_{K302-Q565} intein-CBD fusion protein was purified by Ni-NTA affinity purification via an N-terminal His_{6}-purification tag.

**Figure 5.5** (A) SDS-PAGE of the final product fraction of Ni-NTA-affinity purified LBD_{K302-Q565} intein-CBD fusion protein (lane 1, 61.6 kD) with protein molecular weight standard (lane 2). (B) QTOF LC and MS-spectra of the purified fusion protein (calculated mass after N-terminal methionine cleavage: 61459 Da).

The purification of the ERα LBD_{K302-Q565}-intein-CBD fusion protein by one step affinity column chromatography resulted in 23 mg of pure fusion protein per liter TB expression culture (**Figure 5.5**). The additional purification by size exclusion chromatography prior to the NCL reaction typically did not improve the protein purity significantly (not shown). Intein cleavage and protein thioester ligation to the F-domain peptide could thus subsequently be tested systematically under various conditions.

The hydrolysis of the protein thioester can in theory be circumvented by the addition of the F-domain peptide, since the cysteine side chain thiol-sulfur is a stronger nucleophile than the hydroxyl-oxygen of water engaged in the thioester hydrolysis. Once the product peptide bond has formed after a final S-N acyl shift, the NCL reaction is irreversible. The observed instability of the ERα LBD_{K302-Q565}-thioester might thus be turned into an advantage by yielding a fast transition to the ligation product in the presence of the F-domain peptide. Consequently, the intein-cleavage was supplemented with the F-domain peptide (**Figure 5.6**).
Figure 5.6 (A) SDS-PAGE of the simultaneous intein cleavage and protein ligation (lane 2) as well as intein cleavage in the absence of F-domain peptide (lane 3) and the protein molecular weight standard (lane 1). With the 61.6 kD ERα LBD_{K302-Q565} Mxe intein- CBD fusion protein and the intein cleavage products ERα LBD_{K302-Q565} and CBD- Mxe intein (33.8 kD and 27.9 kD, respectively) and the ligation product ERα LBD-F-pT of 36.8 kD. (B) QTOF-LC and MS-spectra of the simultaneous intein cleavage and native chemical ligation reaction (A, lane 2) with excess peptide (P), intein-CBD by-product (I), residual uncleaved ERα LBD_{K302-Q565} Mxe intein- CBD fusion protein (F) and ligation product (L, calculated mass: 36808 Da).

Simultaneous intein cleavage and NCL by addition of the F-domain peptide to the cleavage reaction at pH 8.0, 4 °C and overnight incubation time, resulted in product formation of the semi-synthetic protein ERα LBD-F-pT (Figure 5.6). Gel image analysis of the SDS-PAGE capture with samples from the one-pot intein cleavage and NCL (Figure 5.6) reveals that 85 to 95 % of the ERα LBD_{K302-Q565} intein-CBD fusion protein is cleaved efficiently in the presence and absence of peptide (Figure 5.6, lane 2 and 3, respectively). Furthermore the F-domain peptide is successfully ligated to yield about 40 % ERα LBD-F-pT product with respect to the available intein-cleaved ERα LBD-thioester (Figure 5.6, lane 2). After successfully introducing the F-domain threonine-phosphorylation at moderate ligation yields, the ligation mixture was further purified by size exclusion chromatography (Figure 5.7). The SDS-PAGE analysis of the ligation product purification by size exclusion reveals the co-elution of ligated and unligated ERα LBD (Figure 5.7, 5). This implies that the protein is intact even after multiple step-purification including a freeze –thaw cycle and extended ligation reaction times at 4 °C, since it is presumably forming heterodimers. On the other hand it illustrates the importance of further improving the yield of protein ligation in order to obtain a homogeneously modified protein sample. The CBD-intein by-product originating from the intein-cleavage reaction was however successfully separated from the ERα- fraction (Figure 5.7, 6).
Figure 5.7 SDS-PAGE of the ERα LBD_{K302-Q365} intein-CBD fusion protein (F, 61.6 kD) purified by Ni-NTA-affinity column chromatography (lane 2) and the reaction mixture of simultaneous intein cleavage and NCL (lane 3) with the 28.7 kD CBD-intein cleavage by-product (lane 3, I), the 33.8 kD unligated ERα LBD (lane 3, E) and the 36.9 kD pT-ERα LBD-F-domain ligation product (lane 3, L).

The ligation mixture was applied to a SEC column, resulting in the elution of the high molecular weight impurities (lane 4), a co-elution of ligated and unligated ERα LBD (lane 5, L and E, respectively) and an elution fraction containing the CBD-intein cleavage by-product (lane 6, I). The fraction containing ligated and unligated ERα LBD (lane 7, same as lane 5; containing L and E) was combined with 2 equivalents of 14-3-3ζ-ΔC (lane 8, Q) resulting in protein mixture (lane 9) with L, E and Q. This protein mixture was applied to the SEC column, resulting in two elution peaks (10 and 11). Peak 1 (lane 10) reveals the co-elution of ligated ERα LBD (L), unligated ERα LBD (E), and 14-3-3ζ-ΔC (Q). Peak 2 (11) contains 14-3-3ζ-ΔC (Q) and unligated ERα LBD (E).

In order to separate the ligated from the unligated protein population (Figure 5.7, 7, same as 5), inhomogeneous ERα-fraction was combined with 2 equivalents of 14-3-3ζ-ΔC (Figure 5.7, 7-9) and subsequently applied to the size exclusion column. In this way a specific co-elution of the ERα LBD-F-pT construct with the 14-3-3 protein was envisioned. Indeed, a co-elution of 14-3-3ζ-ΔC with ERα LBD-F-pT is observed (Figure 5.7, 10), yet this fraction contains the ligated as well as the non-ligated ERα LBD. Strikingly, all of the protein bands have very similar intensities, which may provide more molecular information. Analysis of the band intensities by gel image analysis while taking the respective molecular weight of each protein into account, reveals an approximate 1:1.1:1.7 (ERα LBD-F-pT: ERα LBD : 14-3-3ζ-ΔC) ratio between the protein domains, hinting at the co-elution of an ERα-heterodimer with a 14-3-3ζ-ΔC dimer by first approximation. This hypothesis however needs further elucidation. Interestingly, the major fraction of 14-3-3ζ-ΔC co-elutes together with non-phosphorylated ERα LBD. Also this observation needs in-detail molecular analysis.
In order to understand the mode of interaction between the threonine-phosphorylated ERα construct, 14-3-3 subtypes and the potential role of non-modified ERα, homogeneously modified protein samples need to be produced. To improve the progression of the NCL reaction, a series of altered conditions for the simultaneous intein-cleavage and NCL reaction were explored. Changes in incubation time, temperature, pH and peptide excess were investigated for their effect on ligation efficiency (Table 5.3).

**Table 5.3** SDS-PAGE gel image analysis values of the succession of intein-cleavage (% cleavage) and the relative NCL reaction progression with respect to the available cleavage product (% ligation) at different reaction conditions within 1 h and 60 h reaction time and in the presence of either 5 or 10 equivalents (eq.) of F-domain peptide.

<table>
<thead>
<tr>
<th>Progression of NCL reaction in % after 1 h (60 h) reaction time</th>
<th>peptide</th>
<th>% cleavage</th>
<th>% ligation product</th>
</tr>
</thead>
<tbody>
<tr>
<td>pH 4 °C</td>
<td>5 eq.</td>
<td>33 (82)</td>
<td>88 (30)</td>
</tr>
<tr>
<td></td>
<td>10 eq.</td>
<td>30 (80)</td>
<td>90 (19)</td>
</tr>
<tr>
<td>pH 7.5</td>
<td>5 eq.</td>
<td>34 (91)</td>
<td>84 (21)</td>
</tr>
<tr>
<td></td>
<td>10 eq.</td>
<td>44 (93)</td>
<td>77 (13)</td>
</tr>
<tr>
<td>pH 8.0</td>
<td>5 eq.</td>
<td>35 (77)</td>
<td>100 (32)</td>
</tr>
<tr>
<td></td>
<td>10 eq.</td>
<td>27 (83)</td>
<td>90 (23)</td>
</tr>
<tr>
<td>pH 4 °C</td>
<td>5 eq.</td>
<td>26 (91)</td>
<td>100 (32)</td>
</tr>
<tr>
<td></td>
<td>10 eq.</td>
<td>29 (97)</td>
<td>100 (17)</td>
</tr>
</tbody>
</table>

Taking an error of up to 10% for the image analysis accuracy into account, the systematic analysis of intein-cleavage yield and ligation reaction progression at varied reaction parameters (Table 5.3) revealed an overall low progression of intein-cleavage of around 30 % at short incubation times. The intein cleavage did not appear to be affected significantly by the pH or the temperature of the reaction nor by the amount of available peptide within the short reaction time frame. Furthermore, within the short reaction time frame of 1 hour, the ligation reaction of the available protein thioester proceeded almost to completion at all conditions with slightly lower product formation at pH 7.5 and 20 °C (Table 5.3). One possible explanation for this exception could be that elevated temperatures might increase the oxidative character of the reaction mixture and therefore lower the amount of reduced sulphydryl groups available for the NCL reaction. If this interpretation is correct,
the elevated pH of 8.0 would compensate this effect, since the reaction has been observed to go to 100 % completion at 20 °C and pH 8.0 (Table 5.3). When allowed to proceed for a longer time (Table 5.3, within brackets), the intein-cleavage reached high levels of completion ranging between ~80 % at 4 °C incubation temperature and about 90 % or higher at 20 °C. In contrast to the high yield intein-cleavage, the ligation reaction resides around 20 to 30 %. The reduced efficiency of the NCL reaction at longer incubation times might also hint at a loss of reducing character of the aqueous reaction mixture over time. This hypothesis is supported by the observation of outstandingly low NCL yields at 20 °C and 10 equivalents peptide for both pH values tested.

Overall, the total product formation did not exceed 30 % with respect to the starting amount of ERα LBD<sub>K302-Q565</sub> intein-CBD fusion protein at all conditions. Since the transthioesterification is generally considered the rate limiting step of the NCL reaction, an improved rate of the transthioesterification with the N-terminal cysteine of the F-domain peptide by including a better thiol leaving group was envisioned. Therefore the additional thiol, added to the reaction mixture (MESNA) was replaced by thiophenol, since it has been reported to be an efficient reagent for high yields in NCL reactions. Unfortunately the inclusion of thiophenol led to precipitation of the ERα LBD domain.
5.3 Conclusion

The direct binding behavior of three truncation sequences of the ERα F-domain towards different 14-3-3 subtypes and mutants was characterized in an in-solution FP assay. FP data revealed that threonine-phosphorylation is a strict necessity for the effective protein-peptide interaction. This observation supports the assumption that the cellular ERα-14-3-3 interaction is controlled in a switch-like manner by threonine-phosphorylation. Threonine-phosphorylation of ERα on its penultimate threonine residue has only recently been identified and its resulting regulatory effects are thus still to be elucidated in greater detail. The characterization of the direct protein-peptide interaction in this chapter furthermore emphasizes the overall stabilizing effect of FC on the protein-peptide interaction. Yet, the stabilization of the protein-peptide interaction by FC reveals 14-3-3-variant and peptide-length-specific effects. These significant and consistent differences of the small molecule effect imply the possibility for specifically targeted new drug molecules, stabilizing an ERα-14-3-3 interaction for example in specific estrogen receptor positive breast cancer types which have been described to develop resistance towards endocrine therapy in correlation with tamoxifen-induced high levels of 14-3-3 expression. By therapeutically inducing an intracellular complex-formation between ERα and 14-3-3, ERα transactivation may be inhibited and the roles of 14-3-3 overexpression and ERα transcription factor activity in resistance to tamoxifen therapy may be resolved altogether. This prospect is however based on the assumption that 14-3-3 and ERα interaction results in an inactive protein complex. Whether this assumption describes the true nature of the protein complex still remains to be investigated. Therefore and in order to effectively target the 14-3-3 ERα interaction on a cellular level it is of great importance to understand the molecular determinants and function of this protein-protein interaction, allowing an efficient rational design of molecules with clinical application potential.

The inclusion of a highly conserved sequence region into the pT-peptides was revealed to increase binding affinity between the pT-ERα peptides and the 14-3-3 subtypes significantly (Table 5.2, pT-8mer versus pT-15mer). This observation represents a first indication for specific upstream stabilizing interactions between the ERα sequence and the 14-3-3 domain. Deletions of the 15 most C-terminal residues of the F-domain have previously been shown to have a stimulatory effect on ERα-transactivation. This effect was previously accounted to the in silico predicted β-strand like conserved region at the N-terminus of the 15-mer. The discovery of the penultimate threonine phosphorylation site
and its regulatory effect on 14-3-3 interaction, however, suggests an alternative interpretation of the enhanced ERα transactivation by E2 and tamoxifen upon the deletion of the 15 most C-terminal F-domain residues. In fact, these observations imply the inhibition of ERα transactivation by 14-3-3 complex-formation, which is dependent on the penultimate ERα F-domain pT. The results, described in this chapter furthermore differentiate between the effect of the pT, contained in the short pT-8mer F-domain peptide and the addition of the conserved sequence motif at the N-terminus of the pT-15mer F-domain peptide. The direct interaction characterization by FP reveals that the N-terminal sequence motif adds to the stabilization of the 14-3-3 interaction, yet it does not compensate the loss of the pT, as indicated by the inclusion of the non-phosphorylated T-15mer, which fails to effectively bind 14-3-3 variants. Further elucidation of the interplay between 14-3-3-ERα interaction and ligand effects, which can be speculated to be interconnected by the highly conserved KYYIT-sequence, is of interest for the understanding of this protein-protein cross-talk. A co-crystallization of the semi-synthetic ERα-LBD-F-pT construct with 14-3-3 would ultimately provide insight to the involved residues at molecular and atomic resolution.

In order to achieve such an ERα LBD-F-pT construct via a native chemical ligation approach, expressed protein ligation of the ERα LBD thioester and a modified F-domain peptide were explored and summarized in this chapter. Semi-synthetic ERα LBD-F-pT was successfully produced and included in preliminary co-crystallization screens with 14-3-3σ-ΔC. However, the reproducibility of this protein semi-synthesis turned out to be challenging. The mechanistic details underlying the problems in reproducibility were elucidated and revealed that the ERα protein-thioester is especially labile. As an effective resolution of the encountered fast thioester hydrolysis, a one-pot intein-cleavage and NCL reaction was established. However, the yields of the overall reaction output remained moderate. The slow intein-cleavage was identified as limiting factor at short reaction times whereas the progression of the NCL reaction was identified as limiting factor at long reaction times. An increase of peptide concentration did not resolve the latter limitation. Since the reaction of peptide and protein thioester was observed and saturating peptide concentrations were used, an inhibition of the NCL reaction by disulfide bond formation due to an oxidation of the reaction environment over time provides a feasible explanation of the limited overall turnout of the reaction. In the future, the hypothesis of NCL inhibition by over time oxidation of the reaction mixture should be tested by degassing the solvents prior to mixing the reactants and by incubating the reaction mixture in the absence of oxygen.
As a first approximation, the interaction between ERα LBD-F-pT and 14-3-3 was analyzed by size exclusion chromatography. The co-elution of 14-3-3-dimer together with a mixed ERα LBD-F-pT -ERα LBD dimer was identified, providing a first glance at a possible tetrameric mixed protein complex organization, which would be in line with the findings described in literature.\(^9\) Reported in literature was an enhanced ER dimerization in a yeast two-hybrid assay by a T\(_{594}\)A mutation, presumably by abolishing ERα-14-3-3 interactions. Furthermore immunoprecipitation in the presence of a 14-3-3 peptide-binder competitor presumably also abolished the ER-α-14-3-3 interaction, resulting in enhanced ER dimerization.\(^9\) These observations also allow for the conclusion of the assembly of an ERα-dimer together with a 14-3-3-dimer, resulting in an ααββ-type tetramer-formation and thereby resulting in the indirect inhibition of free ERα dimer formation. Future investigations may elucidate the question whether such a mixed tetramer can also be formed between a 14-3-3-dimer and a homo-dimer of 2 ERα LBD-F-pT molecules and whether this complex formation sterically allows for the interaction of both pT-F-domains with each one binding site of the 14-3-3 dimer. Since this interpretation is speculative and also the co-elution of excess 14-3-3 protein with non-phosphorylated ERα LBD was observed, the described results will need further investigation.
5.4 Experimental section

If not mentioned otherwise, all chemicals and materials were used as purchased from the usual sources and without further purification.

**Peptide synthesis.** The 8mer peptide was ordered from GenScript/USA and used without further purification. The 15mer and 29mer peptides were synthesized on Rink amide resin by automated standard Fmoc-solid phase peptide synthesis (SPPS).\(^{35,36}\) NMP (N-Methyl-2-pyrrolidone) was used as solvent for the entire synthesis protocol. Deprotection of the N-terminal Fmoc-protection group was done in 20 % (v/v) piperidine in NMP. Fmoc-amino acid building blocks were diluted from 200 mM stock solutions to yield a final twofold molar excess of building block over the resin-substitution, reported by the supplier. Coupling reagent HCTU was added at equimolar concentrations and together with 4 equivalents of N,N-Diisopropylethylamine DIPEA. The coupling of each amino acid building block was incubated for 10 minutes. This coupling step was repeated a second time under the same conditions. The remaining free amines after each double coupling step were capped by acetylation in pyridine : NMP : acetic anhydride (1:3:1: ratio). A 5-iodoacetamido-fluorescein building block was used to introduce the fluorescein-labels of the 15mer and the 29mer to the sulfhydryl groups of the N-terminal cysteine residues. Peptides were cleaved from the resin by a 94:2:2:2 ratio of TFA : H\(_2\)O : TIS : EDT (TFA, trifluoroacetic acid; TIS, trisopropylsilane; EDT, 1,2-ethanedithiol) and precipitated in cold diethyl ether during incubation at -20 °C for minimal 1 hour or until precipitation was observed. Successful peptide synthesis was confirmed by analytical LCMS (Table 5.4). High purity of all peptides was assured by purification via preparative reversed phase HPLC with mass detection in adequate gradients from water to acetonitrile in the presence of 0.1 % TFA. Peptides were stored freeze-dried at -20 °C. For the introduction of the phospho-threonine, a Fmoc-Thr(PO(OBzl),OH)-OH building block was used.

**Table 5.4** Summary of peptide names, sequences, theoretical mass in Da and observed mass in Da.

<table>
<thead>
<tr>
<th>Peptide name</th>
<th>Peptide Sequence</th>
<th>Calculated mass (Da)</th>
<th>Observed mass (Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pT-F domain</td>
<td>CHLATAGSTSSHSLQKYITGEAEGFPApTV</td>
<td>3127</td>
<td>3127</td>
</tr>
<tr>
<td>pT-8mer</td>
<td>FAM-A(<em>{588})EGFPApTV(</em>{595})</td>
<td>1229</td>
<td>1229(^a)</td>
</tr>
<tr>
<td>pT-15mer</td>
<td>C(FITC)K(<em>{581})YYITGEAEGFPApTV(</em>{595})</td>
<td>2217</td>
<td>2215</td>
</tr>
<tr>
<td>T-15mer</td>
<td>C(FITC)K(<em>{581})YYITGEAEGFPATV(</em>{595})</td>
<td>2137</td>
<td>2135</td>
</tr>
<tr>
<td>pT-29mer</td>
<td>C(FITC)H(<em>{567})LATAGSTSSHSLQKYITGEAEGFPApTV(</em>{595})</td>
<td>3596</td>
<td>3596</td>
</tr>
</tbody>
</table>

\(^a\)Purchased from GenScript

**Protein expression and purification of 14-3-3 subtypes and mutants.** The different 14-3-3 subtypes and deletion mutants were expressed and purified according to the established 14-3-3 purification procedure from expression vectors, that have been cloned before.\(^{37}\)

**Fluorescence polarization assay (FP).** FP assays were performed in 10 mM Hepes at pH 7.4 with 150 mM NaCl, 0.1 % Tween-20, 0.1 % BSA. All peptides were dissolved in the H\(_2\)O to yield 5 µM stock solutions and subsequently diluted to a final concentration of 20 nM in the assay buffer. 14-3-3 proteins were
Molecular insights into the interaction between 14-3-3 proteins and the ERα F-domain

serially diluted from high µM to nM concentration ranges. Fluorophore polarization was detected on a Tecan Infinite F 500 microplate reader. The resulting data was analyzed by fitting the data with equation $Y = \frac{B_{\text{max}} \times X}{K_D + X} + \text{BG}$. In which $B_{\text{max}}$ is the maximum binding and the polarization level at low concentrations of 14-3-3 protein is included as background (BG).

Protein thioester expression and purification. The overall protocol for protein thioester expression and purification was modified as discussed in the Results and Discussion section of this chapter. The basic procedure of protein thioester expression and purification was done by overnight 25 ml LB medium supplemented with 100 µg/ml pre-cultures from E. coli BL21 (DE) cells freshly transformed with the previously prepared expression vector pHT562. The ERα LBD (M A S W S H P Q F E K G S M G S S H H H H H S S G L V P R G S H M-K302-Q565) is expressed as fusion protein with a C-terminal Mxe intein domain to yield the C-terminal thioester after thiol-induced intein cleavage. A C-terminal chitin binding domain was used for affinity purification on chitin-functionalized resin (New England Biolabs). The entire pre-culture was used to inoculate 2 L TB medium expression culture supplemented with 100µg/ml of ampicillin which was incubated at 37 °C and shaking at 220 rpm until a cell density of $A_{600nm} \sim 0.6$. Protein expression was induced after cooling down the expression medium to 15 °C with 0.5 mM IPTG. After 20 h expression at 15 °C and 220 rpm shaking incubation, cells were harvested by centrifugation for 10 minutes at 7500 rpm and 4 °C. Cell pellets were snap-frozened in liquid N₂ stored at -80°C until further usage. Storage time was kept as short as possible. For the purification of the thioester fusion protein, cells were opened by homogenization (EmulsiFlex C3, Avestin) in Buffer A (20 mM Heps at pH 8.0, 500 mM NaCl, 10 % glycerol, 1 mM EDTA, 10 µM E2, 200 µM PMSF, 0.1 TCEP, 0.5 % Triton X-100, 1 mM MgCl₂, 1 µg/ml DNAse I (EMD Millipore)). Cell-debris was removed by centrifugation for 20000 rpm for 20 minutes at 4 °C. Cell-extract was applied to chitin beads for affinity purification of the fusion protein at 4 °C in Buffer A. Unbound protein was washed off with 10 column volumes buffer B (20 mM Heps at pH 8.0, 500 mM NaCl, 10 % glycerol, 1 mM EDTA, 10 µM E2). On-column intein cleavage was induced at pH 8.5 in Buffer B supplemented with 500 mM MESNA for 16 hours at 4°C. After elution in the same buffer, the cleavage step was repeated. Protein concentration was increased by centrifugation in Amicon® Ultra 10K filter devices and the protein thioester was stored in the elution buffer at -80°C until ligation.

Native chemical ligation. Prior to the native chemical ligation reaction, the F-domain peptide was incubated at 20°C under reducing conditions (2 mM TCEP) to ensure reduced cysteine side chains. For the ligation reactions, protein thioester was incubated with 3 equivalents of the respective H12-peptide overnight at 4 °C and pH 8.0 in Buffer A supplemented with 2 mM TCEP and 200 mM MESNA.
Chapter 5

5.5 Acknowledgements

Anniek den Hamer and Wencke Adriaens were involved in the protein semi-synthesis approach. Lidia Nieto, Anja Schäfer and Anniek den Hamer performed peptide synthesis, labeling and purification of the different length ERα-F-domain peptides. Maria Bartel introduced me to the very reliable FP assay protocol. Christian Haase synthesized the 29mer pT-peptide. Hoang Nguyen and Trang Phan previously cloned the expression vector pH1562.
5.6 References

Protein-protein interaction fine-tuning by estrogen receptor post-translational modification

The estrogen receptor (ER) is a nuclear hormone receptor, mediating the response of cells and tissues to estrogens which are of central interest in the development of primary and secondary reproductive organs. The ER also has regulatory functions in non-apparent gender-affected tissues such as the skeletal and vascular systems. Next to its regulation by its natural hormone ligands, the versatile ER actions are strictly modulated by post-translational modifications (PTMs). The crucial involvement of ER in vital cellular functions is reflected in ER dysregulation in several diseases, amongst which different types of cancer. About 70% of all diagnosed breast cancers depend on functional ER for their progression, which makes ER a highly relevant drug target. Unfortunately, resistance against ER-antagonists presents a common challenge in endocrine therapy. Different pathways to endocrine resistance have emerged, often connected to an altered ER-regulation by PTMs. The detailed molecular basis of the influence of specific PTMs on ER function and structure, however, remained elusive. This is mainly due to the inaccessibility of homogeneously modified ER samples. This dissertation summarizes a number of semi-synthesis approaches towards the introduction of PTMs into the ER designated for the study of their direct effects on ER structure, function and protein-protein interactions in a controlled *in vitro* environment by biophysical techniques. The introduction of a clinically highly relevant tyrosine phosphorylation, which occurs in both ER subtypes and is located in the functionally important C-terminal region of the ligand binding domain of the ER, revealed the potential of a single protein modification to directly regulate the ER-cofactor interaction in a subtype - specific manner. The production of a library of 12 differently labeled and homogeneously phosphorylated ER protein constructs was achieved via an efficient protein semi-synthesis approach. The inclusion of different labels into these protein constructs allowed for the biophysical investigation of the phosphorylation effects on the protein structure and dynamics in the direct vicinity of the modification, including the first protein-NMR data on apo-ER LBD as well as the quantification of the interaction with cofactors by large scale protein interaction screens as well as a small scale direct interaction assays. The results of this investigation are
summarized and discussed in chapter 2, resulting in the hypothesis of a subtype-specific and ligand-independent regulation of ER - activity by a single tyrosine phosphorylation and expanding our general understanding of the molecular determinants of ER activation mechanisms.

Throughout the years a vast amount of knowledge on the function and structure of the two ER subtypes has accumulated. In fact the ERs represent two of the most heavily studied NRs. This results from the clinical relevance and pharmacological potential of these hormone receptors. Increasingly detailed evidence for the ER-regulation in its genomic function as transcription factor as well as a slowly growing non-genomic component of ER-function is emerging. This non-genomic signaling of ER is essentially regulated by PTMs, taking care of the adapted context-specific fine-tuning of ER-function by directly modulating its structural and dynamic character (chapter 2) and by facilitating or abolishing protein-protein interactions with other proteins from the ER genomic and non-genomic signaling networks (e.g. chapter 3). The exact molecular determinants of these PTM-regulated interactions, however, remain largely elusive up to date. The influence of ER PTM on ER-protein interactions, centrally yet not exclusively involved in the regulation of ER non-genomic signaling functions were explored in chapters 3 to 5. The tyrosine phosphorylation-induced interaction between both ER subtypes and c-Src protein tyrosine kinase via the phosphotyrosine interaction domain (SH2 domain) is for the first time quantified in a controlled biochemical environment in chapter 3. Furthermore, the interaction mode between the two proteins is characterized at molecular resolution by NMR spectroscopy. The quantitative characterization of this protein-protein interaction indicates a conserved interaction mode as compared to a high affinity SH2-binding motif and revealed binding affinities, well in-line with commonly reported binding affinities towards other SH2-interaction partners. The discussed interaction studies revealed no binding specificity towards either of the two ER subtypes, hinting at an establishment of this interaction prior to the evolvement of the ER subtypes and emphasizing the potential for the mediation of subtype specific effects via ternary interaction partners. In addition, the presence of agonists or cofactor peptides appeared to leave the ER-SH2 interaction unaffected. This observation emphasizes the dominance of the phosphorylation status in the switch-regulation of this protein-protein interaction and underlines its ligand-independent regulatory function. The ERα hinge region represents another protein-protein interaction surface within the receptors sequence, which is highly decorated, and thus potentially regulated, by PTMs. Chapter 4
Summary

summarizes the characterization of the effects of individual modifications on the hinge region structural characteristics and their effects on the regulation of the specific interaction with ER-regulating interaction partner calmodulin. Opposing effects of hinge region acetylation and phosphorylation have been observed in CD spectroscopy studies. Furthermore, a regulatory effect of hinge region acetylation on the interaction with calmodulin was revealed by interaction studies, which can most certainly be assigned to the specific alteration of the ionic interactions, established at the protein-protein interaction surface. Several approaches towards the semi-synthetic introduction of clinically highly relevant hinge region serine 305 phosphorylation into ERα protein constructs were explored in the same chapter, providing a base for the establishment of strategies for diversely hinge region modified ER protein constructs in the future. The striking co-occurrence between adverse drug effects or drug resistance in endocrine treatment of cancer patients and non-physiological levels of specific ER PTMs, such as the tyrosine phosphorylation characterized in chapter 2 and 3 and the serine phosphorylation addressed in chapter 4, raises the hope for the development of individualized cancer treatment as well as the need for alternative approaches towards a targeted interference with the ER-function in ER-dependent tumors. A recently developed concept for the development of new ER-anagonizing molecules aims for the stabilization of protein complexes which have an inhibitory character on ER-activity. One example of such an interaction is the threonine-phosphorylation induced complex-formation between ERα and several subtypes of the scaffolding-proteins of the 14-3-3 protein family. The recently demonstrated potential for interaction stabilization by a small molecule makes this interaction an attractive target for the development of new anti-cancer drugs. The phosphorylation-dependent interaction between ERα and 14-3-3 proteins is explored in chapter 5 and the stabilization of this interaction by the small molecule fusicoccin is confirmed by these studies. Furthermore, the introduction of the threonine-phosphorylation, which induces the protein-complex formation, into an ERα protein construct was achieved and its limitations evaluated via a semi-synthetic protein ligation approach, providing a detailed experimental base for the future molecular and structural analysis of the ERα F-domain modification.

The molecular basis of ER PTM-effects has not been explored in great detail before, which can mainly be accounted to the inaccessibility of homogenously modified protein samples. As demonstrated in this work, homogeneously modified protein samples carry a great potential to specifically assign structural and functional effects of amino acid side chain modifications within the protein context, which will contribute to the overall understanding
of the complex mechanisms of ER-activity regulation. Next to feeding the pool of knowledge, which is certainly of use for the efficient development of new effective drug molecules or for the rational prediction of the individualized benefits of existing therapies, understanding fundamental effects of amino acid side chain modifications might allow to establish basic principles, which can help interpreting entirely different systems at given analogy. The applied chemical biology approach, employing native chemical ligation facilitated the introduction of protein PTMs and has been proven to be very effective for the simultaneous introduction of read-out labels for biophysical protein characterization to yield detailed insight to the range of influence of a single PTM (chapter 2) and to characterize modification-dependent protein-protein interactions of the receptor which have not been quantified in vitro before (chapter 3). Furthermore, a combination of protein peptide analogues and protein semi-synthesis allows the elucidation of important, yet little characterized protein interactions and reveals the powerful influence of PTMs on intrinsic protein region structural characteristics (chapter 4 and 5). The thorough analysis of challenges, intrinsic to specific ER semi-synthesis approaches provide a base towards the introduction of highly relevant ER modification into the ER protein context (chapter 4 and 5). Altogether the described chemical biology approaches of ER PTM-effects on protein structure, function and protein-protein interaction also provide an example for the future introduction of PTMs into other NRs.

Taken together, the successful and efficient introduction of the C-terminal tyrosine phosphorylation into both ER subtypes together with various site-directed labels for the biophysical analysis of the direct effects of the specific protein modification demonstrate the potential of the established methodology for the elegant investigation of potentially subtype specific PTM-effects, which may also be transferred to other nuclear receptors (chapter 2 and 3). In fact, the established protocols for the partial $^{15}$N-labeling of the C-terminal region of the ER LBD open doors for the detailed investigation of NR functioning and their differential activity regulation by their specific H12 dynamics and its modulation by traditional ligands, novel compounds as well as their interplay with physiologically and clinically relevant PTMs in the highly important C-terminal regulatory NR region. This would presumably enrich our understanding of NR functioning and thus drug-ability fundamentally. In view of the development of novel protein-protein interaction-stabilizers for endocrine intervention, the access to specifically modified NR protein samples is crucial for the elegant and effective exploration of specifically protein-complex targeted drugs, as discussed and worked towards
in chapter 5. The study of PTM-effects in peptide analogues demonstrate the measurable
direct effects of PTMs in their specific environment, which underlines the urgency for the
exploration of regulatory modifications in a protein context in order to understand their
(patho-) physiological function. Introduction of the modifications of interest can be
challenging at times, yet the in-detail investigation of semi-synthesis approaches, discussed
here, will provide a base for the continuation of the introduction of PTMs into the protein
sequence of ER or other NRs and the resulting vast possibilities for NR PTM investigation
and implementation.
### List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ac</td>
<td>acetylation</td>
</tr>
<tr>
<td>AF-1</td>
<td>activation function 1</td>
</tr>
<tr>
<td>AF-2</td>
<td>activation function 2</td>
</tr>
<tr>
<td>CBD</td>
<td>chitin binding domain</td>
</tr>
<tr>
<td>CD</td>
<td>circular dichroism</td>
</tr>
<tr>
<td>CSP</td>
<td>chemical shift perturbation</td>
</tr>
<tr>
<td>DBD</td>
<td>DNA binding domain</td>
</tr>
<tr>
<td>E2</td>
<td>17β-estradiol</td>
</tr>
<tr>
<td>EGF</td>
<td>epidermal GF</td>
</tr>
<tr>
<td>EGFR</td>
<td>EGF receptor</td>
</tr>
<tr>
<td>ERE</td>
<td>estrogen response element</td>
</tr>
<tr>
<td>ER</td>
<td>estrogen receptor</td>
</tr>
<tr>
<td>FC</td>
<td>fusicoccin</td>
</tr>
<tr>
<td>GF</td>
<td>growth factor</td>
</tr>
<tr>
<td>GF-R</td>
<td>GF receptor</td>
</tr>
<tr>
<td>H12</td>
<td>helix 12</td>
</tr>
<tr>
<td>HAT</td>
<td>histone acetyl transferase</td>
</tr>
<tr>
<td>hCaM</td>
<td>human calmodulin</td>
</tr>
<tr>
<td>HDAC</td>
<td>histone deacetylase</td>
</tr>
<tr>
<td>hGR</td>
<td>human glucocorticoid receptor</td>
</tr>
<tr>
<td>hMR</td>
<td>human mineralcorticoid receptor</td>
</tr>
<tr>
<td>hPR</td>
<td>human progesterone receptor</td>
</tr>
<tr>
<td>HRE</td>
<td>hormone response element</td>
</tr>
<tr>
<td>hRXRα</td>
<td>human retinoid X receptor α</td>
</tr>
<tr>
<td>IEC</td>
<td>ion exchange chromatography</td>
</tr>
<tr>
<td>L11</td>
<td>loop 11</td>
</tr>
<tr>
<td>LBD</td>
<td>ligand binding domain</td>
</tr>
<tr>
<td>MAPK</td>
<td>mitogen-activated protein kinase</td>
</tr>
<tr>
<td>MNAR</td>
<td>modulator of non-genomic activity of ER</td>
</tr>
<tr>
<td>NCL</td>
<td>native chemical ligation</td>
</tr>
<tr>
<td>NMR</td>
<td>nuclear magnetic resonance</td>
</tr>
<tr>
<td>NRAM</td>
<td>NR alternative site modulator</td>
</tr>
<tr>
<td>NRTK</td>
<td>non-receptor PTK</td>
</tr>
<tr>
<td>OHT</td>
<td>4-hydroxy-tamoxifen</td>
</tr>
<tr>
<td>pER</td>
<td>phosphorylated ER</td>
</tr>
<tr>
<td>(p)ER</td>
<td>(non-) phosphorylated ER</td>
</tr>
<tr>
<td>PKA</td>
<td>protein kinase A</td>
</tr>
<tr>
<td>PPI</td>
<td>protein protein interaction</td>
</tr>
<tr>
<td>pS</td>
<td>phospho-serine</td>
</tr>
<tr>
<td>pT</td>
<td>phospho-threonine</td>
</tr>
<tr>
<td>PTK</td>
<td>protein tyrosine kinase</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>PTM</td>
<td>post translational modification</td>
</tr>
<tr>
<td>pY</td>
<td>phospho-tyrosine</td>
</tr>
<tr>
<td>RTK</td>
<td>receptor PTK</td>
</tr>
<tr>
<td>SAXS</td>
<td>small-angle X-ray scattering</td>
</tr>
<tr>
<td>SDR</td>
<td>specificity determining region</td>
</tr>
<tr>
<td>SEC</td>
<td>size exclusion chromatography</td>
</tr>
<tr>
<td>SERD</td>
<td>selective ER down-regulator</td>
</tr>
<tr>
<td>SERM</td>
<td>selective ER modulator</td>
</tr>
<tr>
<td>SHR</td>
<td>steroid hormone receptor</td>
</tr>
<tr>
<td>TF</td>
<td>transcription factor</td>
</tr>
<tr>
<td>TF-E</td>
<td>cognate TF-promoter</td>
</tr>
</tbody>
</table>
List of Publications


Tharun, I. M.; Nieto, L.; Balk, M; Haase, C; Wienk, H.; Boelens, R.; Brunsveld, L., Structural and biophysical characterization of the interaction between pY537-ERα and pY488-ERβ with c-Src and v-Src kinase SH2 domain, in preparation.

Nieto, L.; Tharun, I. M.; Canales, A.; Timinez, J.; Wienk, H.; Boelens, R.; Brunsveld, L., CaM binding mode fine-tuning to ERα by hinge region modifications, in preparation.
Inga Maria Tharun was born on 04-01-1984 in Hamburg, Germany, where she finished the Abitur at the Albert-Schweitzer-Gymnasium in 2003. She studied Biochemistry at the Goethe-Universität in Frankfurt/Main and the Universität Witten/Herdecke in Germany, including a research internship in the group of Prof. dr. Marileen Dogterom at the AMOLF in Amsterdam, the Netherlands. She graduated within the group of Prof. Dr. Marina V. Rodnina on the characterization of tRNA^{Asp} kinetics in ribosomal protein translation, receiving the title Diplom-Biochemikerin in 2009. In the same year she started a PhD project at the Eindhoven University of Technology in the Netherlands under the supervision of Prof. dr. ir. Luc Brunsveld on the characterization of protein-protein interaction fine-tuning by estrogen receptor post-translational modifications. The most important results of this work are presented in this dissertation.
Acknowledgements

Finally, I take this opportunity to thank everybody involved in the maturation of this thesis and those, whom have shared this experience with me.

First of all, I express my gratitude to my promoter prof. dr. ir. Luc Brunsveld. Luc, bedankt voor deze periode als PhD in jouw groep, waarin ik zo veel heb mogen leren. Voor zowel mijn wetenschappelijke als ook mijn persoonlijke ontwikkeling gedurende deze jaren wil ik je bedanken voor de vrijheid en jouw vertrouwen, voor jouw respectvolle betrokkenheid en voor jouw open, directe en positieve instelling.

I thank my co-promoter dr. Cristian Ottmann as well as the members of my committee prof. dr. Markus Kaiser, PD dr. Sonja A. Dames, prof. dr. Bert Meijer, dr. Lech-Gustav Milroy, and dr. Maarten Merkx for reading and approving my thesis. Prof. dr. Hilbers, dank u voor uw voorzitterschap. PD dr. Sonja Dames, dear Sonja, thank you very much for the fruitful collaboration, your expertise and input for the phosphotyrosine project and for the personal interest in my future plans. Prof. dr. Markus Kaiser, vielen Dank für die Teilnahme an meinem Komitee. Thank you prof. dr. Bert Meijer, dr. Lech-Gustav Milroy, and dr. Maarten Merkx for the valuable feedback and creative input throughout my PhD project.

I am grateful for having experienced the positive working atmosphere within the Chemical Biology group at the TU Eindhoven. In the beginning of my time as a PhD student in Eindhoven, our group was still very overview-able and I am thankful for the warm welcome by Katja, Dana, Dung, Micha, Sascha, Parisa, Joke and Luc as well as by the members of the Merkx group. I will not forget the apenkoppen-Fridays with Jan and Wencke very soon. Next to the warm welcome in the lab, I have enjoyed the various free-time activities within the Chemical Biology group and built friendships that will certainly last beyond my time at the TU/e. Katja, danke für die Freundschaft, die vielen schönen, lustigen und tiefgründigen Momente. Ich freue mich sehr, dass du meine Paranymphe bist. Wencke, bedankt voor de gezelligheid. Ik zal je missen. Nowadays, the Chemical Biology ‘groepje’ evolved to a large and even more diverse group. I am happy to have been part of this extraordinarily vibrant, creative and driven research environment. I am thankful for all practical help by everybody who has supported me e.g. by carrying heavy stuff while I was carrying my daughter and thereafter.

Remco, Aniek en Wencke bedankt voor de gezelligheid, de goede werk-focus en jullie support tijdens mijn schrijf- en sollicitatie/oriëntatie- fase.
Acknowledgements

Dr. Sabine Möcklinghoff, liebe Sabine, ich bin dir dankbar für die Unterstützung in meiner Anfangszeit in der Chemischen Biologie. Vielen Dank für die wertvolle Vorarbeit zur Introduktion der Tyrosin-Phosphorylierung in die ER LBDs.

Dr. Lidia Nieto, dear Lidia, you expanded the scope of the group with your NMR expertise and working together with you was science at its best to me. I learned a lot from you and I am very grateful for your guidance and enjoyed all lunch break conversations. I am looking forward to stay in contact in the future. Thank you also for joining in the supervision of Anneloes together with Matthew Burton during my pregnancy leave.

I have worked closely together with many more people. Thank you Wencke Adriaens, dr. Matthew Burton, dr. Christian Haase, Andrea Schmidt, dr. Ingrid de Vries-van Leeuwen, Marcel Scheepstra, Maria Bartel, dr. Seppe Leysen and dr. Anja Schäfer. I have learned a lot from you and I am thankful for your input.

Throughout my PhD time, I had the privilege to guide many students. Lisanne Cramer, bedankt voor jouw inzet voor de eerste stappen van het SH2-project tijdens jouw stageproject. Anneloes Oude Vrielink, jij was de eerste studente, die ik tijdens een master project begeleid heb. Ik zal me altijd jouw enthousiasme, creativiteit en doorzettingsvermogen herinneren. Jij bent enorm getalenteerd en ik ben je dankbaar voor je inzet voor het ERα-hinge region project. Ik wens je verder veel plezier en succes met je PhD project. Mark Balk, jij hebt eerst bij mij stage gelopen op het SH2-domain project en bent vervolgens terug gekomen om onder de begeleiding van Ingrid de Vries-van Leeuwen en mij af te studeren. Deze continuïteit was terug te vinden in je manier van werken maar ook in je sociale houding tegenover je omgeving. Het was een plezier om jou te begeleiden en ik ben je dankbaar voor je inzet voor de verschillenden projecten. Ik wens je veel succes met jouw promotieonderzoek. Anniek den Hamer, bij jouw afstudeerproject op het 14-3-3-project was ik pas in de tweede helft betrokken. Jouw uitzonderlijk resultaatgerichte en vrolijke manier van werken laat een blijvende indruk achter. Bedankt voor je inzet voor het 14-3-3 project, dat je mijn paranimfe bent en voor de gezelligheid tijdens mijn schrijffase. Ik wens je dat jouw promotieonderzoek verder zo voorspoedig verloopt.

Furthermore I thank Joost van Dongen together with Ralph Bosmans and Chan Vinh Lam for ensuring the access to the QTOF-LCMS, Hans Damen for his support in ordering chemicals and the delivery of precious samples, Joke and Marjo for their support, and Peggy de Graaf-Heuvelmans for keeping the overview in the LCB. Wencke Adriaens, bedankt voor de introductie in de SPR en de peptidensynthese en voor het ervoor zorgen dat het
Acknowledgements

peptidenlab een prettige werkplek was. Jij was een geweldige leerling in de eiwitexpressie, -zuivering en NCL en ik dank je voor je inzet voor het phosphotyrosine-project.


Buiten de deuren van de universiteit, wil ik de leidsters van onze kinderdagopvang bedanken. Met name de kernleidsters Jolanda, Noortje, Shirley en Angela: Jullie zijn fantastisch en jullie inzet essentieel- dit kan niet vaak genoeg gezegd worden!


Tania, thanks for introducing the word ‘sane’ into my vocabulary. Luc en Janneke, ik ben enorm blij, jullie allebei gedurende mijn promotieonderzoek te hebben leren kennen. Bedankt voor jullie oprechte vriendschap.


Wido, you are the most respectful and caring person I can imagine. I feel deeply honored to have you by my side. Melina, du machst aus uns ein Ganzes. Ich danke euch beiden aus tiefstem Herzen für eure Liebe und Geduld.