

Modulating the nuclear receptor-cofactor interaction : characterization and inhibition

Citation for published version (APA):

Möcklinghoff, S. (2010). *Modulating the nuclear receptor-cofactor interaction : characterization and inhibition*. [Phd Thesis 1 (Research TU/e / Graduation TU/e), Biomedical Engineering]. Technische Universiteit Eindhoven. <https://doi.org/10.6100/IR672740>

DOI:

[10.6100/IR672740](https://doi.org/10.6100/IR672740)

Document status and date:

Published: 01/01/2010

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 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](#)

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.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Modulating the Nuclear Receptor-Cofactor Interaction

Characterization and Inhibition

Modulating the Nuclear Receptor-Cofactor Interaction: Characterization and Inhibition

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 dinsdag 11 mei 2010 om 16.00 uur

door

Sabine Möcklinghoff

geboren te Dortmund

Dit proefschrift is goedgekeurd door de promotor:
prof.dr.ir. L. Brunsveld

This research has been financially supported by Merck & Co., Bayer-Schering Pharma and Merck Serono

Cover Design: Sabine Möcklinghoff and Krumbacher Media & Art
Printing: CPI, Wöhrmann Print Service, Zutphen

A catalogue record is available from the Eindhoven University of Technology Library

ISBN: 978-90-386-2210-1

Table of Contents

Chapter 1	General Introduction	1
1.1	The Nuclear Receptor Superfamily	2
1.1.1	The Estrogen Receptor	5
1.1.2	The Androgen Receptor	6
1.2	Targeting the Nuclear Receptor – Cofactor Interaction	7
1.2.1	Inhibitors of the Nuclear Receptor – Cofactor Interaction	10
1.3	Post-translational Modifications on Estrogen Receptors	13
1.4	Chemical Biology Approaches on Nuclear Receptors	16
1.5	Aim and Outline of the Thesis	19
1.6	References	20
Chapter 2	On-Bead Peptide Screening for Modulators of the Androgen Receptor – Cofactor Interaction	27
2.1	Introduction	28
2.2	Design and Synthesis of a Combinatorial Peptide Library	29
2.3	On-Bead Screening and Biochemical Peptide Hit Validation	31
2.4	Conclusion	41
2.5	Experimental Section	42
2.6	References	47
Chapter 3	Miniproteins as Modulators of the Androgen Receptor – Cofactor Interaction	49
3.1	Introduction	50
3.2	Computational Design and Synthesis of the Miniproteins	51
3.3	Androgen Receptor Binding Studies of the Miniproteins	56
3.4	Co-Crystallization Approach of Miniproteins in Complex with the Androgen Receptor Ligand Binding Domain	65
3.5	Estrogen Receptor Binding Studies of the Miniproteins	67
3.6	Co-Crystallization Approach of the Miniprotein Apa-3 in Complex with the Estrogen Receptor β LBD	75
3.7	Conclusion	76
3.8	Experimental Section	77

3.9	References	85
Chapter 4	Protein Semi-Synthesis and Evaluation of the Phosphorylated Estrogen Receptor β Ligand Binding Domain	89
4.1	Introduction	90
4.2	Protein Semi-Synthesis of the Phosphorylated ER β LBD	92
4.3	Structural Analysis using <i>Circular Dichroism</i> Spectroscopy	99
4.4	Binding Studies using a Cofactor Recruitment On-Chip Assay	103
4.5	Binding Studies using a Cofactor Recruitment FRET Assay	106
4.6	Co-Crystallization Studies of the Phosphorylated ER β LBD in Complex with Estradiol and Cofactor Peptide	111
4.7	Binding Studies using Surface Plasmon Resonance (SPR)	117
4.8	Conclusion	120
4.9	Experimental Section	121
4.10	References	130
Chapter 5	Exploring the Role of the Phosphorylated Estrogen Receptor α Ligand Binding Domain via Semi-Synthesis	133
5.1	Introduction	134
5.2	Protein Semi-Synthesis of the Phosphorylated ER α LBD	135
5.3	Structural Analysis using <i>Circular Dichroism</i> (CD) Spectroscopy	139
5.4	Binding Studies using a Cofactor Recruitment On-Chip Assay	140
5.5	Binding Studies using a Cofactor Recruitment FRET Assay	143
5.6	Co-Crystallization Studies of the ER α LBD with Estradiol and Cofactor Peptide	147
5.7	Binding Studies using Surface Plasmon Resonance (SPR)	151
5.8	Conclusion	152
5.9	Experimental Section	153
5.10	References	160
Chapter 6	Fragment-Based Estrogen Receptor Ligands	163
6.1	Introduction	164
6.2	Identification, Design and Fragment-Based Synthesis of Novel Estrogen Receptor Ligands	165

6.3	<i>In vitro</i> Structure – Activity Relationship of Novel Estrogen Receptor Ligands	167
6.4	Transactivation Efficiency of the Novel Estrogen Receptor Ligands <i>in vivo</i>	172
6.5	Co-Crystallization of Novel Estrogen Receptor β Ligands	175
6.6	Conclusion	181
6.7	Experimental Section	183
6.8	References	188
	Summary	191
	Curriculum Vitae	196
	Acknowledgement	197

List of Abbreviations

Å	Angstroem	EGF	Epidermal growth factor
AA	Amino acid	EPL	Expressed protein ligation
Ac	Acetyl	Eq.	Equation
AF-1	Activation function 1	equiv	Equivalent
AP-1	Activator protein 1	ER	Estrogen Receptor
APC	Allophycocyanin	ERE	Estrogen response element
AR	Androgen Receptor	ESI	Electrospray ionization
ATP	Adenosine triphosphate	<i>et al.</i>	<i>et alii</i>
BF	Binding function	Eu	Europium
BSA	Bovine serum albumin	FITC	Fluoresceinisothiocyanate
Bza	(2-benzothiazolyl)alanine	Fmoc	9-fluorenylmethyl-oxycarbonyl
calcd	calculated	FP	Fluorescence Polarization
CBD	Chitin binding domain	FPLC	Fast protein liquid chromatography
CBI	Coactivator binding inhibitors	FRET	Förster Resonance energy transfer
CD	circular dichroism	GFR	Growth factor receptor
CDK	Cyclin-dependent kinase	GPCR	G-protein-coupled receptor
CEX	cation exchange chromatography	GST	Glutathione-S-Transferase
CFP	Cyan fluorescent protein	H12	Helix 12
Da	Dalton	HBTU	2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium
DAB	3,3'-diaminobenzidine	HDAC	Histone deacetylases
DBD	DNA binding domain	HEPES	4-(2-hydroxyethyl)-piperazine-1-ethane-sulfonic acid
DCM	Dichloromethan	HOBt	1-hydroxybenzotriazole
DES	Diethylstilbestrol	HPLC	High pressure liquid chromatography
DHT	5 α -dihydrotestosterone	HRT	Hormone replacement therapy
DIC	Diisopropylcarbodiimide	HSP	Heat shock protein
DIPEA	N,N-diisopropyl-ethylamine	HT	4-hydroxytamoxifen
DMF	Dimethylformamide	HTS	High-throughput screening
DMSO	Dimethylsulfoxide	IMAC	ion-metal affinity chromatography
DNA	desoxyribonucleic acid	IPTG	Isopropyl- β -D-thiogalactopyranoside
dNTP	Deoxynucleotide triphosphate	ITC	Isothermal titration calorimetry
DTT	1,4-dithiothreitol	K _D	Dissociation constant
E2	17 β -estradiol	K _I	Inhibition constant
E6-AP	E6-associated protein	LAGE	Light-activated gene expression
EC ₅₀	Half maximal effective concentration	LB	Lysogeny broth
EDS	N-ethyl-N'-dimethylaminopropyl) carbodiimide hydrochloride	LBD	Ligand binding domain
EDT	1,2-ethanedithiol	LCMS	Liquid chromatography/mass spectroscopy
EDTA	Ethylenediamino-N,N,N',N'-tetraacetic acid	MALDI	Matrix-assisted laser desorption/ionization

MAPK	Mitogen-activated protein kinase	Strep	Streptavidin
MESNA	2-mercaptoethanesulfonic acid	syn	synthetic
MNAR	Modulator of non-genomic activity of ER	TB	Terrific broth
MW	Molecular weight	TBST	Tris-buffered saline with Tween 20
MWCO	Molecular weight cut off	TCEP	Tris(2-carboxyethyl)phosphine
Nap	(2-naphthyl)alanine	TES	Triethylsilyl
NCL	Native chemical ligation	TFA	Trifluoroacetic acid
NCoA	NR coactivator	TFE	Trifluoroethanol
NHS	N-hydroxysuccinimide	ThioPh	(2-thio)phenylalanine
NFkB	Nuclear factor kappa B	THIQ	Tetrahydroisoquinoline
NMR	Nuclear magnetic resonance	TIF	Transcription intermediary factor
NR	Nuclear Receptor	TIS	Triisopropylsilane
OBOC	One-bead-one-compound	T _M	Melting temperature
OD	Optical density	TOF	Time of flight
PAGE	Polyacrylamide-gel electrophoresis	TR	Thyroid receptor
PBS	Phosphate-buffered saline	Ub	Ubiquitylation
PCR	Polymerase chain reaction	UV	Ultra violet
PDB	Protein data bank	VIS	visible
PEG	Polyethylene glycol	YFP	Yellow fluorescent protein
<i>Pfu</i>	<i>Pyrococcus furiosus</i>		
PhCl	(4-chloro)phenylalanine		
PKA	Protein kinase A		
pmp	4-phosphonomethyl-L-phenylalanine		
PMSF	Phenylmethylsulfonylfluorid		
PTM	Post-translational modification		
Qdot	Quantum dot		
RAL	Raloxifen		
RAR	Retinoic acid receptor		
RP	Reverse-phase		
rpm	Runs per minute		
RXR	Retinoid X factor		
SAR	Structure-activity relationship		
SDS	Sodium dodecyl sulphate		
SEC	Size exclusion chromatography		
SERM	selective ER modulator		
SH2	Src homology-2		
SH3	Src homology-3		
SMRT	Silencing mediator of RAR/TR		
SPPS	Solid-phase peptide synthesis		
SPR	Surface Plasmon resonance		
SRC	Steroid receptor coactivator		

Chapter 1

General Introduction

Abstract: This chapter contains a brief introduction to some important topics in this thesis. First, the nuclear receptor family is introduced with the focus on the estrogen and the androgen receptors. Then, the concept of inhibiting the nuclear receptor – coactivator interaction is explained and the contribution of post-translational modifications on nuclear receptor function is demonstrated. A short overview of the use and efficiency of chemical biology studies on nuclear receptors are given. Finally, the scope of the thesis is outlined.

1.1 The Nuclear Receptor Superfamily

Nuclear Receptors (NRs) are members of a large superfamily of evolutionary related multi-domain transcription factors typically under the control of small lipophilic molecules that easily penetrate biological membranes and allow the NR to specifically regulate the expression of target genes^[1]. Thereby, NRs influence a variety of physiological processes, including the regulation of reproductive systems by steroid hormones, the control of development by thyroid and retinoid hormones and the regulation of bile acid and cholesterol biosynthesis^[2]. Apart from the normal physiology, NRs have been identified to play a role in many pathological processes, such as cancer, diabetes, rheumatoid arthritis, asthma or hormone-resistance syndromes^[3].

Despite the functional diversity exhibited by this class of transcription factors, they share a remarkable structural and functional similarity^[4]. Most NRs are comprised of single polypeptide chains that can be separated into four to five modular domains^[5] (Figure 1). Although there is undoubtedly interplay between the domains, each domain has a separate crucial signal transduction function. The activation function 1 (AF-1), which is located on the poorly defined N-terminal A/B domain, is important for the ligand-independent constitutive transcriptional activity of the NR^[6]. The central DNA binding domain (DBD) is based on two highly conserved cysteine-rich zinc finger motifs, which bind, in response to a ligand, specifically DNA sequences in the promoters of target genes^[7]. The hinge region is a mobile linker between the DBD and the C-terminal ligand binding domain (LBD) facilitating the DBD and the LBD to adopt different conformations^[4]. The overall architecture of the LBD is comprised of 12 anti-parallel α -helices and is well-conserved between the various family members^[8]. Helices 1-11 form a structure comprising a ligand-binding pocket that diverges sufficiently among the NRs to guarantee selective ligand recognition. The flexible C-terminal helix 12 (H12), which is commonly referred to as the ligand-induced activation function (AF-2) is crucially involved in transcriptional coregulator interactions^[9]. A further structural distinct region within the LBD is a dimerization surface, which mediates the interaction with the partner LBDs^[10].

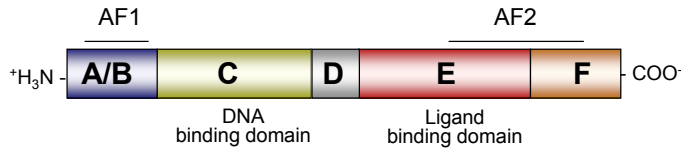


Figure 1: General modular structure of nuclear receptors. The N-terminal A/B domain, including the activation function 1, is important for the constitutive transcription activity. The central DNA binding domain is followed by a hinge region (D) and the ligand binding domain that includes the activation function 2, crucial for the ligand-dependent cofactor interaction. A few nuclear receptors feature an additional C-terminal F domain, whose exact function is still poorly understood^[11].

Historically, NRs have been classified into three subgroups^[12]. In the absence of ligand, type I receptors, including steroid receptors like the androgen receptor (AR) and the estrogen receptor (ER), are located in the cytosol of target cells where the LBD is often associated with heat shock proteins (HSP)^[13] (Figure 2). In this transcriptional inactive conformation, H12 is distended away from the body of the receptor^[8b]. Binding of an agonistic ligand, such as the natural hormone estradiol for the ER or testosterone for the AR, induces a significant conformational change within the LBD. As a result, the NR dissociates from the heat shock proteins, causes non-genomic effects or translocates into the nucleus and site-specifically binds DNA as NR homodimer^[14]. In this NR conformation, the ligand is trapped in the core of the LBD by a conformational shift of H12 against the surface of the LBD that partially obscures the pocket^[14a, 15]. This repositioning of H12 completes the formation of a favourable shallow, hydrophobic groove on the surface of the LBD formed by H3, H4 and H12. This hydrophobic cleft is recognized by transcriptional coactivator proteins^[16] that bind it via a so-called α -helical LXXLL motif (where L is a leucine and X is any amino acid)^[17]. The coactivator peptide is held in place through interactions of its three hydrophobic residues and hydrophobic groove constituents but also by hydrogen bonds between peptide backbone atoms and two well-conserved residues, a lysine at the C-terminus of H3 and a glutamate in H12, that together form a charge clamp to further stabilize the NR – coactivator interaction^[18] (Figure 3,4). Genetic and biochemical data have revealed a plethora of coactivator proteins that mediate NR function, including the p160 transcription intermediary factor/NR-coactivator/steroid receptor coactivator (TIF/NCoA/SRC) family of proteins^[19]. These coactivators often have multiple LXXLL motifs, termed NR boxes, of differing affinity that can exhibit either cooperative or non-cooperative binding. Beside the ability to bind to the NR, the coactivators feature transcriptional-enhancing properties, including histone

acetylation and chromatin remodelling, as well as the recruitment of additional transcription factors and other elements of the transcription machinery^[20].

Type II receptors, including the thyroid receptor (TR) and the retinoic acid receptor (RAR), contrast with type I receptors in terms of localization and DNA-binding orientation. These receptors are retained in the nucleus and bind the DNA response element constitutively as heterodimers with the retinoid X factor (RXR)^[20]. Further, these receptors have the ability to stably interact with the corepressors NCoR and SMRT via a LXXXIXXXI/L motif^[21]. This motif adopts a three turn helix that docks to a hydrophobic groove on the surface of the LBD that overlaps with the upper part of the AF-2. Thus, H12 is displaced in the unligated state to expose the corepressor binding site^[22]. The complex of DNA-bound NR and corepressor recruits transcriptional complexes that contain specific histone deacetylases (HDACs), resulting in target gene repression^[23]. In most cases, agonist binding promotes complex allosteric effects that lead to an alternative positioning of H12, disrupting the hydrophobic groove and releasing the corepressor^[24] (Figure 2).

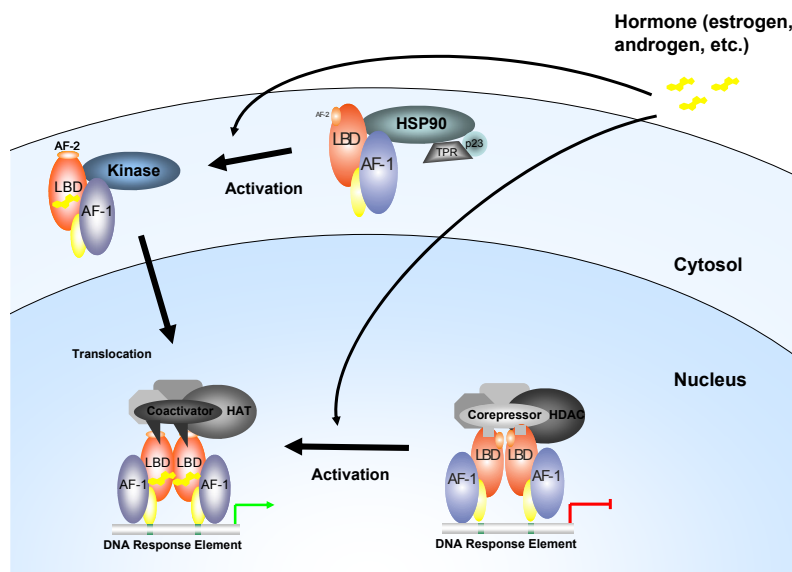


Figure 2: Schematic, simplified representation of nuclear receptor (NR) action. Interaction of a ligand with the NR LBD enhances the phosphorylation of the NR^[25] and exerts non-genomic effects or results in nuclear translocation and promoter binding. Some NRs are bound to the regulatory regions of target genes as a corepressor complex in the absence of agonists. Ligand binding releases this complex and causes the recruitment of coactivator complexes resulting in target gene transcription.

As the ligand-regulated NR – coactivator interaction is the crucial element in the mediation of transcriptional activation, disruption of this protein-protein interaction has become a suitable target in drug discovery in the recent years. Up to recently, the modulation of the function of NRs concentrated mainly on ligands with agonistic, antagonistic or partial (ant)agonistic properties that bind in the binding pocket of the LBD, also occupied by the natural ligands^[10a, 16b, 26]. Interestingly, there are also NRs for which, up to now, no ligand, neither of natural nor of synthetic origin, has been found. This third type of 48 known human NRs, also called orphan receptors do not seem to be amenable to modulation by classical ligands^[27]. Due to the large number of nuclear receptors, here it will be focused predominantly on the estrogen receptor and the androgen receptor.

1.1.1 The Estrogen Receptor

The action of estrogens and their analogous in regulating gene expression is mediated mainly by the two subtypes ER α ^[28] and ER β ^[29] that feature unique tissue distribution patterns. While ER α is predominantly present in the uterus and mammary gland and is mainly involved in reproductive events, ER β is the more generally expressed ER and seems to be relevant in the central nervous system, bone, lung, cardiovascular system, ovary, testis, urogenital tract, kidney, and colon^[30]. Although the LBDs of both subtypes share only 47% homology^[30-31], their ligand binding cavities differ by only two amino acids (ER α Leu₃₈₄ → ER β Met₃₃₆; ER α Met₄₂₁ → ER β Ile₃₇₃)^[32].

Both ER α and ER β are mainly regulated by the natural female sex hormone, the estrogen 17 β -estradiol (E₂)^[33]. Certain types of breast cancer and postmenopausal diseases have been found to be estrogen dependent^[34]. Synthetic estrogens like diethylstilbestrol (DES)^[35] bind to the ER at the same position as E₂ and cause the same conformational change in the LBD (Figure 3a). Anti-estrogens^[36] and selective ER modulators (SERMs)^[37] are often used to (partially) inhibit the estrogen function in cancer cells. Basic side-chain extensions of these ligands prevent productive positioning of H12 against the ligand binding pocket and induce a distinct conformation of H12 at a position that overlaps with the coactivator-binding groove of the receptor and often leads to the recruitment of corepressors^[16b]. As a result, the α -helical LXXLL motifs of coactivator peptides can not be recognized anymore resulting in the transcriptional inhibition of the NR (Figure 3b). SERMs, such as raloxifen (RAL)^[38] and 4-hydroxytamoxifen (HT)^[39] that display tissue-specific or –selective agonist/antagonist pharmacology have been proven to be very beneficial for the treatment of breast cancer and the regulation of hormone functioning^[26d, 40]. Nevertheless, predictability and control over

tissue specificity of SERMs is difficult, in particular as their physiological effects are sometimes also altered with time^[26h]. A typical example of this drug resistance is the transition of HT from an antagonistic profile into an agonistic profile over a period of 1-2 years^[41].

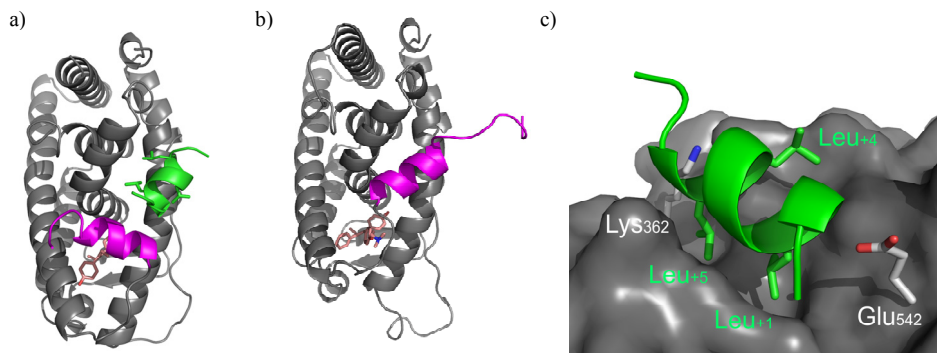


Figure 3: (a) Crystal structure of the wildtype ER α LBD bound with the estrogen DES and a coactivator peptide (PDB: 3ERD^[26c]); (b) the ER α LBD bound to the anti-estrogen 4-hydroxytamoxifen (PDB: 3ERT^[26c]). Ligands are shown as stick model; Helix 12 is shown in purple and the coactivator peptide is shown in green. (c) Detail of the helical LXXLL peptide bound to the hydrophobic groove on the ER α LBD surface between the charge clamp residues Lys₃₆₂ and Glu₅₄₂.

1.1.2 The Androgen Receptor

The Androgen Receptor (AR) belongs to the type I steroid receptors and is responsible for the development of the male sexual characteristics, but also plays a role in the occurrence of certain types of prostate cancer and types of breast cancer^[42]. AR function is regulated by the binding of androgens, including testosterone and its reduced form 5 α -dihydrotestosterone (DHT), which initiates sequential conformational changes of the receptor LBD^[43] (Figure 4a). The resulting formation of a hydrophobic groove on the surface of the LBD is crucial for both the amino/carboxyl-terminal (N/C) interaction of the AR and the coactivator recruitment during transcriptional activation^[44]. Unlike other steroid receptors, the hydrophobic groove of the agonist-bound AR is longer, deeper, narrower and smoother and features a relatively low affinity for most, but not all LXXLL motifs^[18a, 18b]. Instead, this groove primarily accommodates the AR N-terminus^[45] or AR specific coregulators (ARA70, ARA54, etc.)^[45a, 46] that feature unique recognition motifs, in which the leucines of the typical LXXLL motif are replaced by bulkier aromatic phenylalanines (FXXLF) or tryptophans (WXXLF). The backbone of a bound AR-specific motif forms hydrogen bonds with both charge clamp residues (K₇₂₀ and E₈₉₇) instead of only one hydrogen bond as with LXXLL motifs^[18a, 47]

(Figure 4b). Further, it was shown that the AR hydrophobic groove binds the more compact LXXLL motifs via subsidiary contacts with flanking negative charged residues^[47].

The N/C interaction is an important feature for the stabilization of H12, slowing down the ligand dissociation from the LBD^[48]. However, especially upon DNA binding, depending on the type of promoter, this interaction loses importance^[49]. Although binding of the AR specific coregulators enhances the transcriptional activity, it is assumed that gene-transcription in the AR is mainly regulated by the interaction of conventional coactivators and corepressors with the activation function 1 (AF-1) domain^[50]. Certain mutations in the AR as well as coactivator overexpression increase AR functional activity via binding of LXXLL-containing cofactors to the AF-2 and are thought to be significant contributors to the development of prostate cancer^[18b, 44, 51]. Treatment with AR antagonists, like bicalutamide^[52] and hydroxyflutamide^[53] does not initiate the N/C interaction and results in the recruitment of corepressors. However, with time (a few month to years), cellular modifications including point mutations in the AR often convert these ligands into potent AR agonists resulting in the resistance to their antagonistic action^[54]. Additionally, up-regulation of the AR and coactivators as well as different phosphorylation patterns on the AR influence the agonist/antagonist behaviour^[54d, 55].

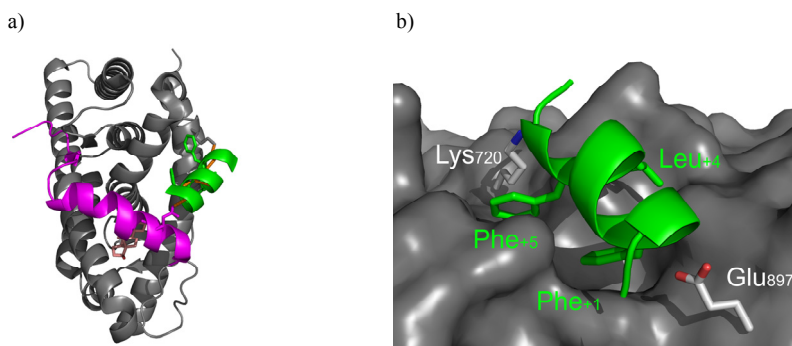


Figure 4: (a) Crystal structure of the wildtype AR LBD bound to the androgen DHT and a FXXLF peptide motif (PDB: 1T7R)^[18a]. The ligand is shown as stick model; Helix 12 is shown in purple and the coactivator peptide is shown in green. (b) Detail of the helical FXXLF peptide bound to the hydrophobic groove on the AR LBD surface between the charge clamp residues Lys₇₂₀ and Glu₈₉₇.

1.2 Targeting the Nuclear Receptor – Cofactor Interaction

The limitations of the current NR ligands generate a continuous quest for new ligands with optimized profiles. The requirements of the effects of these ligands on the level of the organism are usually quite clear; the translation of these requirements to the level of the conformation of the protein and structure of the ligand is, however, usually not evident. This

is mainly due to the large plethora of protein – protein interactions between NRs and their many cofactors^[26h, 56]. There are several hundreds of cofactor proteins known, but both the molecular rules for the formation of the NR – cofactor complexes and the physiological meaning of many of these are often still a mystery^[57]. Additionally, ligand binding also influences other NR – protein interactions (*e.g.*, with the nuclear factor kappa B; NFκB^[58] or the activator protein-1; AP-1^[59]) located at topically different positions on the NR. The identification of compounds that can inhibit the interaction of the NR with its cofactor by directly binding to the NR interface would facilitate the selective inhibition of a special NR – coactivator interaction without or only partially influencing other NR – protein interactions. Such coactivator binding inhibitors (CBIs) would be important chemical tools in the biological elucidation of this protein – protein interaction on the molecular level and would have the promise of yielding new types of NR antagonists (Figure 5). Additionally, CBIs might provide an entry for the modulation of orphan NRs for which, up to now, no classical ligand has been found.

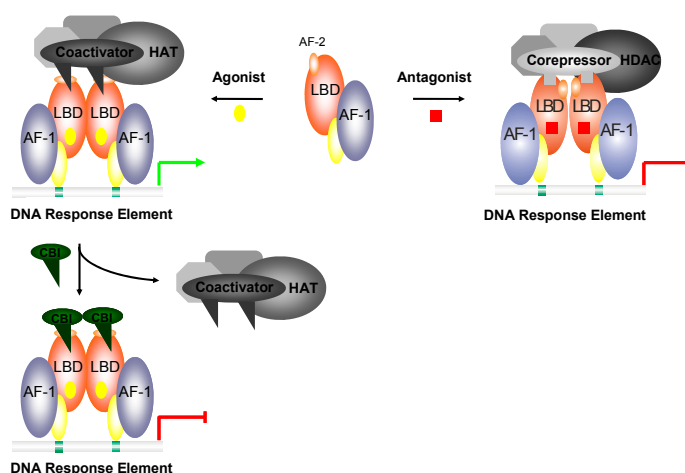


Figure 5: Schematic, simplified representation of the modulation of nuclear receptor functioning via classical agonists (yellow) resulting in the recruitment of coactivators featuring, among others, histone acetyltransferase (HAT) functionalities and via classical antagonists (red) resulting in the recruitment of corepressors featuring, among others, histone deacetylase (HDAC) functionalities. A new approach is the modulation of nuclear receptor functioning via CBIs (green) targeting either agonist-liganded nuclear receptors or orphan NR, resulting in the displacement of the coactivator complex.

It has been already proven that direct inhibition of a NR – coactivator interaction is a valid principle to selectively antagonize NR functioning. The ER – coactivator interaction could be directly inhibited by coactivators in mammalian cells, resulting in the inhibition of

ER-mediated gene transcription^[41a]. Further, specific peptides containing LXXLL motifs, identified via phage display, are able to bind to the ER – cofactor interface in the cell and feature the inhibition of estradiol-mediated transcriptional activity in a cell type dependent manner^[24, 41a, 60]. However, linkage of two of the identified LXXLL peptides even increased the antagonistic effect^[24]. In further studies, a 100-fold molar excess of short LXXLL-containing peptides from natural coactivators was needed to displace 50% of a full-length coactivator from ER α or ER β ^[61]. Many coactivators feature more than one LXXLL motif and bind to NR dimers in a multivalent fashion, leading to strong binding. Beside the AF-2, the interaction with the AF-1 contributes to the overall affinity. Inhibition of only the AF-2 via CBIs might not be sufficient for completely antagonizing NR functioning. In order to establish the direct inhibition of the NR – coactivator interaction as a *bona fide* pharmaceutical target, this issue requires addressing. In particular, for the AR the AF-1 plays a significant role in transactivation. Different peptides containing FXX(L/F/M/Y)F, (F/W)XXL(F/W), FXXLY or WXXVW^[18a, 62] motifs, identified via phage display against full-length AR, were shown to be capable of suppressing the AR N/C interaction and a selected set of these peptides was actually capable of suppressing AR transactivation in the presence of any ligand. These peptides showed high affinity and selectivity for the AR. This high affinity and a different functional inhibition mode could be the reasons that these peptides are able to overcome the occupancy of the AF-2 with the N-terminal FXXLF motif and also effectively antagonize the AR. Some of these peptides showed almost complete inhibition of AR-mediated gene expression when coexpressed in cells, proving the possibility to effectively antagonize the AR in cells with a CBI^[63].

CBIs could further be used to overcome the undesired resistance of traditional antagonists after a certain period of time^[64]. Interestingly, LXXLL-containing peptides that were shown to inhibit estradiol-mediated transcriptional activity via the ER were observed not to be able to block ER α transcription mediated by 4-OH-tamoxifen (HT)^[41a]. Peptides isolated against HT-activated ER α therefore generally did not contain the LXXLL motif. Expression of these peptides in cells blocked the partial agonistic activity of HT-liganded ER α up to 90%, while having no or only a minimal effect on estradiol-mediated transcription^[41a]. These and other peptides^[65] generated against HT-bound ER α or ER β are known to interact with a region of the LBD that is not affected by the distinct conformational effects induced by receptor agonists and antagonists^[66]. This unique interaction surface may play a role in the sensitivity of ERs to coregulators in the presence of antagonists like HT. Thus, this is another nuclear receptor – cofactor interaction that offers the possibility for the development of

pharmaceuticals antagonizing the partial agonistic activity of, for example, 4-hydroxy-tamoxifen.

1.2.1 Inhibitors of the Nuclear Receptor – Cofactor Interaction

A successful inhibitor of the NR – coactivator interaction (CBI) requires a potent and selective profile on the NR. Chemically different types of CBIs have already been developed in the past such as peptides and small molecules. The identification of potent peptide binders for the NR – cofactor interface (*e.g.* for both ER α and ER β and AR) was typically performed via phage display as demonstrated before. Together with other studies, the phage display-derived peptides allowed the identification of amino acid characteristics at different positions in the peptide required to induce NR selectivity^[18a, 47, 63, 67]. In combination with mutational analysis, phage display constitutes a powerful method for developing peptide sequences that can specifically bind to a given NR. Additionally, this methodology is useful for the identification of unknown NR-binding proteins and for the analysis of different NR surface conformations generated by different hormone ligands. Importantly, the nature of the protein, the LBD alone or the complete protein can have a crucial influence on the outcome of the phage display screening and on the functionality of the peptides.

In addition to the composition of the peptide sequence, the α -helical fold of the CBI is crucial for high affinity binding to the hydrophobic groove of the NR. Cyclization of peptides is an established strategy to stabilize short peptides in an α -helical conformation. Concomitantly, different types of cyclic peptides have been explored as CBIs. The introduction of macrolactam bridges^[68] (**1**, Figure 6) as well as side-chain – side-chain disulfide bridges, using a D,L-dicysteine^[69] (**2**, Figure 6) motif, into LXXLL-containing peptides led to the identification of α -helical CBIs with high affinities and selectivities for the ER. Peptide screening in the presence of different hormones^[68c] enabled the possibility to generate CBIs that selectively target different conformations of the same NR, which is of importance for the design of inhibitors for specific cell types or for NR mutants featuring different conformations. Introduction of specific non-natural amino acids in the cyclic peptides positively influenced the binding affinity^[70]. Cyclic peptides with other thiol-containing amino acids, such as homocysteine and penicillamine were also reported^[70] (**3**, Figure 6). Studies on peptides that are not cyclic and feature only one Cys residue have revealed a strong affinity to the ERs. It turned out that these peptides bind to Cys residues in the LBD via disulfide bridges^[70].

Recently, the focus on identifying novel CBIs has shifted to the development of small molecule scaffolds that possess pharmaceutical potential due to their low molecular weight, improved bioavailability and potential for high binding selectivity of these compounds. Thus, small molecule CBIs are envisaged to provide a new entry in NR antagonism. Based on computational approaches and crystallographic information, chemical scaffolds were designed featuring pendant substituents mimicking the Leu side chains of the LXXLL motif or Phe side chains in the FXXLF motif. With respect to hydrophobic interactions with the ER α hydrophobic groove and ionic interactions with the charge clamp, compounds were identified that effectively and selectively inhibited coactivator recruitment. One of these compounds, guanylhydrazone, effectively inhibited coactivator recruitment with IC₅₀ values in the low micromolar range^[71] (**4**, Figure 6). Katzenellenbogen *et al.* designed a series of guanylhydrazone derivatives and confirmed that these compounds are true CBIs and not conventional antagonists^[72]. High-throughput screening (HTS) using fluorescence polarization and time-resolved fluorescence resonance energy transfer (TR-FRET)^[73] led to the identification of trisubstituted pyrimidine scaffolds in which the three leucines of the LXXLL motif are sufficiently mimicked by the alkyl moieties^[74] (**5**, Figure 6). Due to the intrinsic torsional flexibility of pyrimidine cycles, the hydrophobic groups attached seemed to be mobile enough to adopt an optimized position and, therefore, showed a good affinity for the hydrophobic groove of the ER. Structure – activity relationship (SAR) of a series of synthesized pyrimidine derivatives facilitate a refined pharmacophore model for CBIs of this compound class^[75]. Hexasubstituted benzenes with alternating hydrophobic groups were designed to improve the positioning of the leucine-mimicking side-chains. Additionally, hydrophilic groups mimic the amphipathic nature of natural coactivator peptides and allowed the interaction with the exposed solvent^[76]. Incorporation of multiple bulky substituents in these pyrimidine cores (more than two aromatic rings), mimicking the AR-specific FXXLF motif, resulted in complete AR-selective CBIs^[77] (**6**, Figure 6). These compounds have also been shown to inhibit coactivator binding to the T877A AR mutant that is partial responsible for antagonist resistance. The design of a series of bicycle[2.2.2]octane scaffolds that are close structural mimics of the two key leucine residues of the SRC LHRL motif resulted in CBIs with modest potency to block the ER α – coactivator interaction^[78].

An α -helical proteomimetic approach, described by Hamilton *et al.*, provides an alternative to small molecular scaffolds. Bi- and triaryl scaffolds replicate the α -helical rotation of the peptide backbone and display aliphatic or benzylic substituents in the positions of the hydrophobic side chains of the LXXLL motif^[79] (**7**, Figure 6). Katzenellenbogen *et al.*

designed bipolar bis-4,4'-oxyphenyl scaffolds that address both the substitution pattern of the hydrophobic core and the electronic interactions of the charge clamp^[80] (**8**, Figure 6).

A crystal structure of the ER β LBD revealed a folded alternative disposition of 4-hydroxytamoxifen sitting on the coactivator recognition surface of the ER, overlapping with the corresponding coactivator peptide^[81]. Although its physiological relevance needs to be proven, this is the first crystallographic characterization of a small molecule bound to the hydrophobic cleft of a nuclear receptor, which can give new insights in the effort of developing small molecules as potent CBIs. Regarding the thyroid receptor (TR), Fletterick *et al.* have reported selective β -aminoketones^[82] that target the TR β by covalently binding to the hydrophobic groove via thiol nucleophiles displayed on the NR surface in analogy to the previously described peptides^[70]. Fletterick *et al.* could further identify another class of small-molecule CBIs that neither bind to the ligand binding pocket (binding function 1; BF-1) nor to the coactivator binding surface (BF-2) on the AR. These compounds, including non-steroidal anti-inflammatory drugs^[83] and the thyroid hormone 3,3',5-triiodothyroacetic acid^[84], block AR activity via binding to a previously unknown regulatory surface cleft termed BF-3^[85]. This interaction site allosterically influences coregulator association with the NR AF-2 domain and is a known target of mutations in prostate cancer and the androgen insensitivity syndrome.

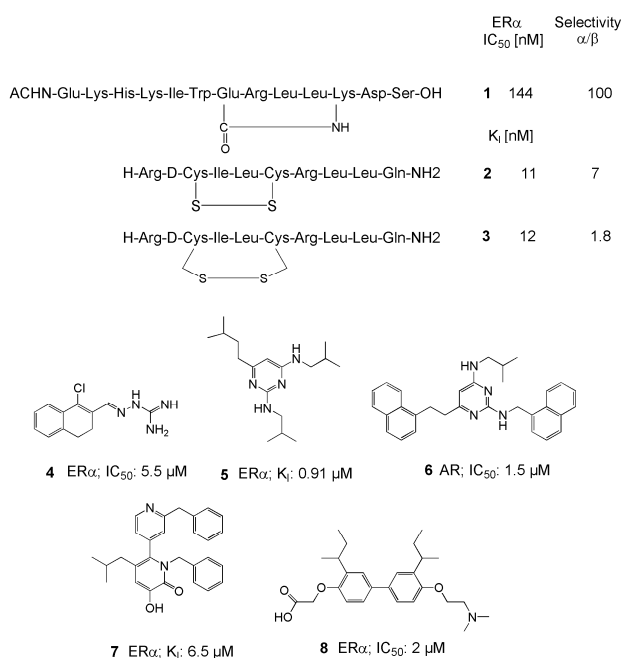


Figure 6: Chemical structures, binding potencies and selectivities of selected cyclic peptides and small-molecule CBIs for ER α , ER β and AR.

The so far developed CBIs indicate the potential of targeting the NR – coactivator interaction and, considering the limited amount of optimization work performed on these compounds thus far, their moderate affinity holds great promise for significant improvement. As such, these compounds provide a very good basis for optimization, proof-of-concept and selectivity studies. Independent of the ultimate biological validity of these compounds, these recent studies show that new and optimized CBIs can be developed with which the concept of inhibition of the NR – cofactor interaction can be investigated in more detail and be possibly validated for many NR.

1.3 Post-translational Modifications on Estrogen Receptors

The transcriptional activity of nuclear receptors (NR) is mainly governed by ligand binding and the recruitment of coactivators, but many processes of NR function are additionally regulated by post-translational modifications (PTM)^[86]. A prominent example is the phosphorylation^[87], but also other types of modification, such as acetylation^[88], ubiquitylation^[89], sumoylation^[90] and palmitoylation^[91], have been demonstrated. These modifications affect, among others, the sensitivity of hormone response, receptor expression and stability, subcellular localization, dimerization, DNA binding and protein-protein interactions. Crosstalk between NR-mediated and other signal-transduction pathways is undoubtedly also an important aspect of NR action. Several non-genomic signalling pathways modify the receptor post-transcriptional, alter its function in a ligand-independent manner and thus play major roles in the resistance against antiestrogens and the partial agonistic behaviour of SERMs.

NR phosphorylation has been studied for much longer than other PTMs. NRs are phosphoproteins and phosphorylation seems to affect most, if not all, NRs. As the regulation of function is receptor specific, the ER is used as an example to present the different effects of phosphorylation on NRs. Due to the difficult availability of post-translational modified proteins for *in vitro* analysis, the roles of phosphorylations have been studied mainly in cell lines or using mutants to prevent or mimic phosphorylation. The majority of the remarkable number of phosphorylation sites lies within the amino-terminal A/B region, but several sites are also located in the DBD and the LBD of both the ER α ^[87a] and the ER β ^[92] (Figure 7). Most of the modified residues are serines surrounded by prolines that correspond to consensus sites for proline-dependent kinases including the cyclin-dependent kinases (CDK-7)^[93] and the MAPK family^[94]. Phosphorylation of Ser₁₁₈ and less so Ser_{104/106} and Thr₃₁₁ by these kinases often facilitates the subcellular localization^[95] and enhances/reduces the recruitment of

coactivators or components of the transcription machinery^[93-94, 96] and therefore cooperates with the ligand to regulate transcription activation (Figure 8). Phosphorylation of Ser₁₆₇ by kinases, including casein kinase II^[97], or Ser₂₃₆ by the protein kinase A (PKA)^[98] supports the transactivation activity of ER α through the regulation of DNA binding^[99] or dissociation^[98, 100]. More pronounced effects of ER phosphorylation on the regulation of receptor activity are noted when specific cell signalling pathways are activated. For example, ER α can also be activated in a hormone independent manner by rapid signalling through epidermal growth factor (EGF) activated MAPK or AKT pathways via phosphorylation of Ser₁₁₈^[101] or Ser₁₆₇^[102] (Figure 8). Thus, deregulation of NR phosphorylation in certain diseases or cancers may lead to apparently ligand-independent activities or resistance against ER antagonists. The recruitment of coactivators and RNA polymerase by ER α phosphorylation of Ser₃₀₅, *e.g.* is enhanced following tamoxifen treatment^[103]. Additionally, Ser₃₀₅ phosphorylation prevents Lys₃₀₃ acetylation and sensitizes ER α to ligand stimulation^[104]. The role of phosphorylation of Tyr₅₃₇, localized in the ER α LBD, has been more controversial^[105]. Mutation of this site alters estradiol binding kinetics and some amino acid substitutions promote hormone-independent activation^[106]. Clearly, some of the changes in function are not simply a consequence of phosphorylation or a lack of phosphorylation. Therefore, other methods than mutagenesis are needed to address the exact role of Tyr₅₃₇ phosphorylation in ER α function. However, Tyr₅₃₇ is a prominent example to point out, how NRs can achieve the autoregulation of their transcription activity. The ER has been shown to form a pre-complex with the membrane-associated protein MNAR (modulator of non-genomic activity of ER)^[107] and/or with the AR^[108]. Androgen, estrogen or EGF are assumed to promote the association of this ER α complex with the SH2 domain of the Src tyrosine kinase via the phosphorylated Tyr₅₃₇ resulting in the activation of Src and downstream MAPK pathways that in turn regulate ER function (Figure 8).

In addition to the interaction of the ER with membrane-associated proteins, non-genomic signalling can occur via direct membrane-associated ER. Marino *et al.* discovered that S-palmitoylation of a cysteine in the ER LBD (conserved among all hormone nuclear receptors) is required for membrane templating, caveolin-1 binding and subsequent non-genomic receptor activity^[109] (Figure 7). While estradiol induces the dissociation of ER α from the plasma membrane to allow the interaction with *e.g.* MNAR, hormone binding to the ER β increases its association to the membrane^[91]. Thereby, palmitoylation raise the presumption that PTMs might be able to individually regulate ER α and ER β activity under specific conditions. Indeed, both ERs can efficiently recruit the coactivator SRC-1, but only for ER β

this process is enhanced through AF-1 phosphorylation^[106b, 110] (Figure 7). Moreover, the phosphorylation-mediated ER β /SRC-1 complex is ineffectively disrupted by the SERM tamoxifen, suggesting that ER β activation could occur in cancer featuring therapeutic resistance^[106b, 111]. In contrast to ER α , a unique modification site could be identified in the mouse ER β (mER β). Phosphorylation of Ser₈₀ enhances receptor degradation^[112]. Ser₈₀ is also a site for O-GlcNAc modification^[112-113]. The two competing modifications likely regulate the stability of ER β . Further, the proteasome-mediated receptor degradation in a Ser₉₄ and S₁₀₆ phosphorylation-dependent manner, could only be found in ER β ^[114] and very recently, it was reported that hER β can be activated via G-protein-coupled receptor (GPCR) induced phosphorylation of Ser₈₇^[115] (Figure 8). Similar to ER α , a tyrosine phosphorylation site in the ER β LBD has been identified that facilitates constitutive receptor activity. Auricchio *et al.* reported that ER β can also form a complex with the AR for the ligand-dependent activation of Src^[116] (Figure 7). However, further investigation is needed to address the exact role of tyrosine phosphorylation in ER β .

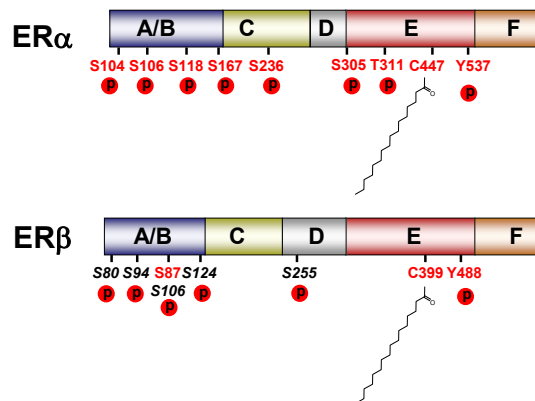


Figure 7: Schematically representation of the estrogen receptor α and β and the localization of some of their proven or potential post-translational modification sites (red for human ER and black for mouse ER)^[87a, 92]. P: phosphorylation.

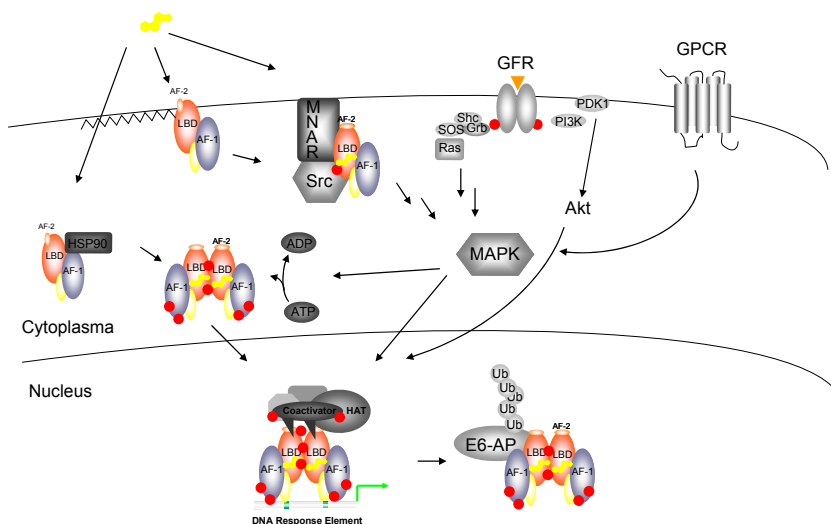


Figure 8: Schematic model illustrating the signalling networks involved in ER action. Beside the classical genomic transcriptional activity, membrane-associated palmitoylated ERs also bind hormones and initiate the interaction with Src and *e.g.* MNAR, activating downstream MAPK pathways. Alternatively, rapid signalling through growth factors and associated growth factor receptors (GFR) or GPCRs can activate kinase pathways that enhance phosphorylation of the ER (illustrated by red spheres) and converge upon and activate target genes or facilitate receptor degradation by ubiquitylation (Ub).

1.4 Chemical Biology Approaches on Nuclear Receptors

Biologically, NRs are of profound importance due to their ability to directly influence a variety of physiological processes and their role in the development of many diseases. Chemically, NRs are of interest as they offer the basis to develop chemical tools for modifying and directly controlling NR function, structure and interactions with other biomolecules on the molecular level. The combination of biology and chemistry, termed chemical biology, uses chemical methods and tools to selectively elucidate and manipulate certain fundamental biological processes on the molecular level^[117].

While the described CBIs facilitate selective cofactor-based control over certain NR functions, chemical biology additionally offers the possibility to selectively control NR-based cellular process via chemically synthesized hormone sensors^[117]. Koh *et al.* developed novel NR LBD variants with unique ligand specificities that respond to complementary ligands with unprecedented selectivity to provide better control over biological systems^[118] (Figure 9a). The incorporation of certain point mutations and polar group exchanges in the LBD facilitates the NR to bind complementary to tailored ligand structures. As the Δ NR can not interact with natural ligands and the synthetic ligand can not be recruited by endogenous NRs, target gene

expression is solely under the control of the functionally orthogonal Δ NR – synthetic ligand pair without cross-reactivity^[119]. Thereby, Δ NR – synthetic ligand pairs demonstrate an attractive target to rescue drug resistance due to NR mutations^[120]. Alternatively, the NR LBD could be tailored to the structural requirements of a non-natural ligand through selection and directed evolution in yeast^[121] with using additional rational design^[122] or via step-wise *in vitro* ligand – NR coevolution^[123].

It was reported for a multitude of NRs that light-activated gene expression (LAGE) provides a possibility to enforce both spatial and temporal control over gene expression on a subtissue specific scale by using photo-caged, inherently bioavailable, small NR ligands^[124]. Whereas these systems enable transient gene expression or repression, permanent activation or silencing of gene expression can also be achieved using photo-activated tamoxifen-dependent recombination^[125]. Toshima *et al.* took a step further and used light-activated ligands for the selective degradation of the ER α leading to fragmentation under visible light irradiation^[126].

Similar developments in the hormone-based control of target protein degradation have been done by Crews *et al.* via chemical inducers of dimerization^[89b, 127]. These heterobifunctional small molecules consist of a ligand that recruits E3 ubiquitin ligase linked to either dihydroxytestosterone (DHT) that targets the AR or linked to estradiol that targets the ER (Figure 9b). Treatment of breast or prostate cancer cells with the corresponding synthetic ligands resulted in the inhibition of cell proliferation by post-translational ubiquitylation and degradation of the AR or ER.

Strategies for HTS of interactions between various hormones and drugs with a certain NR are crucial for accelerating the understanding of NR biology and pharmacology. In this regard, Umezawa *et al.* used the intracellular folding pattern of the AR to distinguish between AR agonists and antagonists in living cells^[128]. Herein, an AR LBD – coactivator peptide fusion protein is flanked by the fluorescent proteins CFP and YFP. Agonist induced conformational changes of the AR LBD recruit the coactivator peptide to the hydrophobic groove and thus brings the two fluorophores in close proximity resulting in FRET. This intramolecular rearrangements can be inhibited by antagonists and allow the rapid detection of ligand binding processes without the need of target gene transcription. Gambhir *et al.* were able to combine the ability to tailor NR LBDs with new ligand binding specificities to characterize ER ligands for their activity *in vivo*^[129]. Several synthetic antagonists, but not estradiol, were able to bind to the mutant ER α G521T fused to split fragments of a

Renilla/firefly luciferase reporter. The ligand activity could easily be studied via bioluminescence due to the split protein complementation (Figure 9c).

An additional method to use chemical NR ligands as sensors to induce the split protein complementation could be demonstrated by Liu *et al.*^[130]. This intein splicing method was based on the conditional protein splice-switch established by Muir *et al.*^[131] through which the autocatalytic excision of inteins is achieved by induced pairing of a split-peptide. Liu *et al.* reported the direct evolution of an intein-based molecular switch in which intein-splicing in yeast cells was made dependent on tamoxifen-binding to a split intein – ER LBD fusion protein. This method was used to post-translational modulate protein function *in vivo*^[132]. Moreover, a related intein-based technology might be a good opportunity to investigate post-translational NR modifications *in vitro*. Expressed protein ligation (EPL) is a semi-synthetic technique in which a recombinant protein thioester, generated by thiolysis of an intein-fusion protein, is reacted with a synthetic or recombinant peptide with an N-terminal cysteine to produce a native peptide bond (Figure 9d). EPL has been used extensively for the incorporation of biophysical probes, unnatural amino acids or post-translational modifications in proteins^[133]. As this methodology has been shown to be applicable to many systems^[133b], expanding EPL on the development of selectively post-translational modified NRs is promising to provide new insights in NR function on the molecular level, including protein-protein interactions and structural information, and their contribution in genomic as well as non-genomic cellular activity.

All presented chemical biology methods have been proven fruitful for the selective control of specific functions of certain NRs without affecting endogenous signalling pathways. New strategies are expected to enable an even better understanding of the complex NR pathways through the continuing development of novel hormone- or coactivator-analogues. Expanding the use of chemical tools to develop selective and homogeneously post-translationally modified NRs will be valuable to understand the complex interplay between NR function and many other signalling pathways.

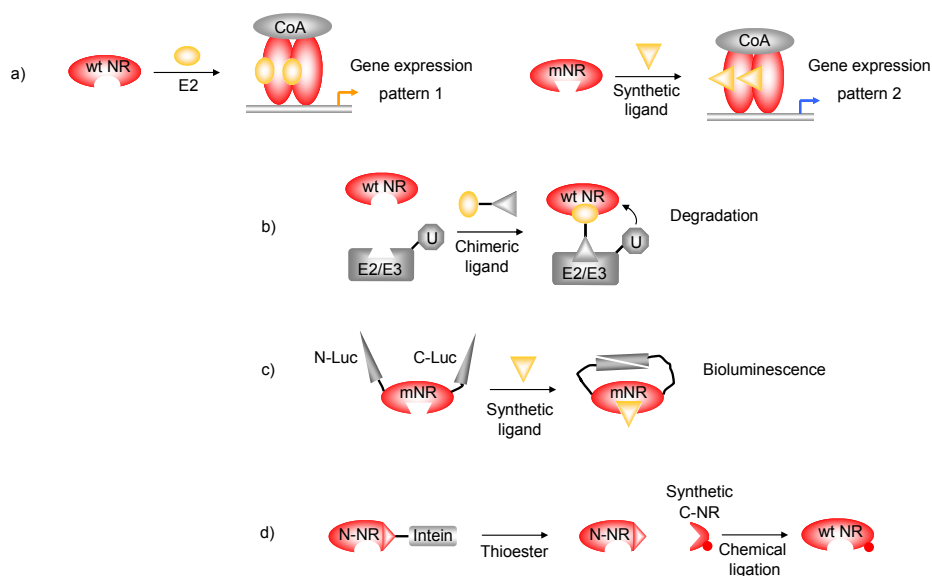


Figure 9: Schematically model of the recent advancements in selectively controlling and investigating NR function via chemical biology techniques. (a) Concept of alternative gene patterning with synthetic ligand – mutant NR (mNR) orthogonality. (b) Target NR degradation based on a chimeric ligand. (c) Evolved ligand – induced split luciferase complementation. (d) EPL as a strategy to investigate ER function. Thioester generation of an N-terminal NR fragment via intein-based thiolysis and native chemical ligation of a synthetic modified NR C-terminus with an N-terminal cysteine. wt: wild type, U: Ubiquitination.

1.5 Aim and Outline of the Thesis

It is increasingly evident that nuclear receptor (NR) action is regulated by a dynamic network of interacting proteins in which molecular recognition events, conformational changes and post-translational modifications have an important role in linking the proper response to a specific signal. The NR – cofactor interaction has been recognized as pharmacological intervention due to their involvement in certain diseases. The use of synthetic scaffolds has been shown as a powerful tool to selectively study and understand molecular NR events *in vitro*. Moreover, NR assembly by using chemical biology technologies is assumed to apply for the study of specific molecular recognition events in the NR – coactivator interactions. The aim of this thesis is the investigation of the factors that are responsible for the selective molecular recognition of NRs by their many cofactors using a combination of molecular biology, chemical synthesis and biophysical characterization.

The development and elucidation of a rapid screening technique for novel peptide-based AR coactivator binding inhibitors (CBIs), which provides the opportunity to combine HTS of large peptide libraries with the possibility to incorporate non-natural amino acids, is described

in **chapter 2**. Additionally to the peptide sequence of the coactivator recognition motif, the secondary fold into a short amphipathic α -helix is essential for high affinity receptor binding. In **chapter 3**, computational design and chemical synthesis were applied to generate helix stabilized miniproteins bearing the coactivator recognition motif for the AR. Biophysical techniques were developed to identify highly potent AR binders and to determine the contribution of helix stabilization and length for selective AR binding.

Post-translational modifications can significantly contribute to the NR – coactivator interaction. The generation of selective and homogenously phosphorylated NRs is necessary to enable a better understanding of the molecular correlations between phosphorylation and NR interaction with other proteins. **Chapter 4** describes the generation of the ER β LBD, site-specifically phosphorylated on Tyr₄₈₈, via intein-based protein semi-synthesis. Different cofactor binding experiments were performed to confirm the correct fold of the semi-synthetic protein and to study the influence of Tyr₄₈₈ phosphorylation on cofactor binding. Crystallization studies of a phosphorylated ER β LBD were planned to provide new structural insights in ER regulation by Tyr₄₈₈ phosphorylation. Expanding on this study, a similar semi-synthesis approach and the characterization of the ER α LBD phosphorylated on the corresponding Tyr₅₃₇ was described in **chapter 5**.

A plethora of synthetic ligands for NRs were synthesized with differentiated effects on cofactor binding. Fragmentation of ligands into discrete functional groups would simplify the analysis of ligand binding and the search of different elements required for selective coactivator recruitment. **Chapter 6** addresses the fragment-based design of novel ER agonists based on a simple tetrahydroisoquinoline scaffold. Biophysical, cell-biological and structural studies helped to further understand the mechanism of ER β affinity and selectivity and provide the entry to develop new agonists that selective induce the recruitment coactivators by the ER β .

1.6 References

- [1] J. Sonoda, L. Pei, R. M. Evans, *FEBS Lett* **2008**, 582, 2.
- [2] N. Novac, T. Heinzel, *Curr Drug Targets Inflamm Allergy* **2004**, 3, 335.
- [3] a)K. Wang, Y. J. Wan, *Exp Biol Med (Maywood)* **2008**, 233, 496; b)W. Shao, M. Brown, *Breast Cancer Res* **2004**, 6, 39; c)L. Di Croce, S. Okret, S. Kersten, J. A. Gustafsson, M. Parker, W. Wahli, M. Beato, *EMBO J* **1999**, 18, 6201; d)M. K. Hansen, T. M. Connolly, *Curr Opin Investig Drugs* **2008**, 9, 247; e)J. L. Staudinger, K. Lichti, *Mol Pharm* **2008**, 5, 17.
- [4] P. Germain, B. Staels, C. Dacquet, M. Spedding, V. Laudet, *Pharmacol Rev* **2006**, 58, 685.
- [5] a)V. Giguere, S. M. Hollenberg, M. G. Rosenfeld, R. M. Evans, *Cell* **1986**, 46, 645; b)A. Krust, S. Green, P. Argos, V. Kumar, P. Walter, J. M. Bornert, P. Chambon, *EMBO J* **1986**, 5, 891.
- [6] A. Warnmark, E. Treuter, A. P. Wright, J. A. Gustafsson, *Mol Endocrinol* **2003**, 17, 1901.
- [7] a)B. F. Luisi, W. X. Xu, Z. Otwinowski, L. P. Freedman, K. R. Yamamoto, P. B. Sigler, *Nature* **1991**, 352, 497; b)J. W. Schwabe, L. Chapman, J. T. Finch, D. Rhodes, *Cell* **1993**, 75, 567.

- [8] a)D. Moras, H. Gronemeyer, *Curr Opin Cell Biol* **1998**, *10*, 384; b)W. Bourguet, M. Ruff, P. Chambon, H. Gronemeyer, D. Moras, *Nature* **1995**, *375*, 377; c)R. T. Nolte, G. B. Wisely, S. Westin, J. E. Cobb, M. H. Lambert, R. Kurokawa, M. G. Rosenfeld, T. M. Willson, C. K. Glass, M. V. Milburn, *Nature* **1998**, *395*, 137; d)R. K. Bledsoe, V. G. Montana, T. B. Stanley, C. J. Delves, C. J. Apolito, D. D. McKee, T. G. Consler, D. J. Parks, E. L. Stewart, T. M. Willson, M. H. Lambert, J. T. Moore, K. H. Pearce, H. E. Xu, *Cell* **2002**, *110*, 93.
- [9] a)W. Bourguet, P. Germain, H. Gronemeyer, *Trends Pharmacol Sci* **2000**, *21*, 381; b)Y. Li, M. H. Lambert, H. E. Xu, *Structure* **2003**, *11*, 741.
- [10] a)W. Bourguet, V. Vivat, J. M. Wurtz, P. Chambon, H. Gronemeyer, D. Moras, *Mol Cell* **2000**, *5*, 289; b)R. T. Gampe, Jr., V. G. Montana, M. H. Lambert, A. B. Miller, R. K. Bledsoe, M. V. Milburn, S. A. Kliewer, T. M. Willson, H. E. Xu, *Mol Cell* **2000**, *5*, 545.
- [11] a)M. Ruff, M. Gangloff, J. M. Wurtz, D. Moras, *Breast Cancer Res* **2000**, *2*, 353; b)B. Farboud, M. L. Privalsky, *Mol Endocrinol* **2004**, *18*, 2839.
- [12] H. Gronemeyer, J. A. Gustafsson, V. Laudet, *Nat Rev Drug Discov* **2004**, *3*, 950.
- [13] a)D. Picard, *Trends Endocrinol Metab* **2006**, *17*, 229; b)W. B. Pratt, D. O. Toft, *Endocr Rev* **1997**, *18*, 306.
- [14] a)W. Gao, C. E. Bohl, J. T. Dalton, *Chem Rev* **2005**, *105*, 3352; b)A. C. Pike, A. M. Brzozowski, R. E. Hubbard, *J Steroid Biochem Mol Biol* **2000**, *74*, 261.
- [15] R. L. Wagner, J. W. Apriletti, M. E. McGrath, B. L. West, J. D. Baxter, R. J. Fletterick, *Nature* **1995**, *378*, 690.
- [16] a)J. M. Wurtz, W. Bourguet, J. P. Renaud, V. Vivat, P. Chambon, D. Moras, H. Gronemeyer, *Nat Struct Biol* **1996**, *3*, 206; b)A. M. Brzozowski, A. C. Pike, Z. Dauter, R. E. Hubbard, T. Bonn, O. Engstrom, L. Ohman, G. L. Greene, J. A. Gustafsson, M. Carlquist, *Nature* **1997**, *389*, 753; c)B. D. Darimont, R. L. Wagner, J. W. Apriletti, M. R. Stallcup, P. J. Kushner, J. D. Baxter, R. J. Fletterick, K. R. Yamamoto, *Genes Dev* **1998**, *12*, 3343.
- [17] D. M. Heery, E. Kalkhoven, S. Hoare, M. G. Parker, *Nature* **1997**, *387*, 733.
- [18] a)E. Hur, S. J. Pfaff, E. S. Payne, H. Gron, B. M. Buehrer, R. J. Fletterick, *PLoS Biol* **2004**, *2*, E274; b)B. He, R. T. Gampe, Jr., A. J. Kole, A. T. Hnat, T. B. Stanley, G. An, E. L. Stewart, R. I. Kalman, J. T. Minges, E. M. Wilson, *Mol Cell* **2004**, *16*, 425; c)A. Warnmark, E. Treuter, J. A. Gustafsson, R. E. Hubbard, A. M. Brzozowski, A. C. Pike, *J Biol Chem* **2002**, *277*, 21862.
- [19] a)H. Chen, R. J. Lin, R. L. Schiltz, D. Chakravarti, A. Nash, L. Nagy, M. L. Privalsky, Y. Nakatani, R. M. Evans, *Cell* **1997**, *90*, 569; b)S. A. Onate, S. Y. Tsai, M. J. Tsai, B. W. O'Malley, *Science* **1995**, *270*, 1354; c)Y. Kamei, L. Xu, T. Heinzel, J. Torchia, R. Kurokawa, B. Gloss, S. C. Lin, R. A. Heyman, D. W. Rose, C. K. Glass, M. G. Rosenfeld, *Cell* **1996**, *85*, 403; d)H. Hong, K. Kohli, A. Trivedi, D. L. Johnson, M. R. Stallcup, *Proc Natl Acad Sci U S A* **1996**, *93*, 4948; e)J. J. Voegel, M. J. Heine, C. Zechel, P. Chambon, H. Gronemeyer, *EMBO J* **1996**, *15*, 3667; f)J. Torchia, D. W. Rose, J. Inostroza, Y. Kamei, S. Westin, C. K. Glass, M. G. Rosenfeld, *Nature* **1997**, *387*, 677; g)S. L. Anzick, J. Kononen, R. L. Walker, D. O. Azorsa, M. M. Tanner, X. Y. Guan, G. Sauter, O. P. Kallioniemi, J. M. Trent, P. S. Meltzer, *Science* **1997**, *277*, 965; h)H. Li, P. J. Gomes, J. D. Chen, *Proc Natl Acad Sci U S A* **1997**, *94*, 8479.
- [20] N. J. McKenna, B. W. O'Malley, *Endocrinology* **2002**, *143*, 2461.
- [21] a)H. J. Huang, J. D. Norris, D. P. McDonnell, *Mol Endocrinol* **2002**, *16*, 1778; b)J. D. Chen, R. M. Evans, *Nature* **1995**, *377*, 454; c)A. J. Horlein, A. M. Naar, T. Heinzel, J. Torchia, B. Gloss, R. Kurokawa, A. Ryan, Y. Kamei, M. Soderstrom, C. K. Glass, et al., *Nature* **1995**, *377*, 397; d)P. Ordentlich, M. Downes, W. Xie, A. Genin, N. B. Spinner, R. M. Evans, *Proc Natl Acad Sci U S A* **1999**, *96*, 2639; e)E. J. Park, D. J. Schroen, M. Yang, H. Li, L. Li, J. D. Chen, *Proc Natl Acad Sci U S A* **1999**, *96*, 3519; f)Y. Shang, X. Hu, J. DiRenzo, M. A. Lazar, M. Brown, *Cell* **2000**, *103*, 843.
- [22] a)A. Marimuthu, W. Feng, T. Tagami, H. Nguyen, J. L. Jameson, R. J. Fletterick, J. D. Baxter, B. L. West, *Mol Endocrinol* **2002**, *16*, 271; b)J. Zhang, X. Hu, M. A. Lazar, *Mol Cell Biol* **1999**, *19*, 6448; c)L. J. Burke, M. Downes, V. Laudet, G. E. Muscat, *Mol Endocrinol* **1998**, *12*, 248; d)X. Hu, M. A. Lazar, *Nature* **1999**, *402*, 93; e)L. Nagy, H. Y. Kao, J. D. Love, C. Li, E. Banayo, J. T. Gooch, V. Krishna, K. Chatterjee, R. M. Evans, J. W. Schwabe, *Genes Dev* **1999**, *13*, 3209; f)V. Perissi, L. M. Staszewski, E. M. McInerney, R. Kurokawa, A. Krones, D. W. Rose, M. H. Lambert, M. V. Milburn, C. K. Glass, M. G. Rosenfeld, *Genes Dev* **1999**, *13*, 3198; g)H. E. Xu, T. B. Stanley, V. G. Montana, M. H. Lambert, B. G. Shearer, J. E. Cobb, D. D. McKee, C. M. Galardi, K. D. Plunket, R. T. Nolte, D. J. Parks, J. T. Moore, S. A. Kliewer, T. M. Willson, J. B. Stimmel, *Nature* **2002**, *415*, 813.
- [23] H. B. Hartman, J. Yu, T. Alenghat, T. Ishizuka, M. A. Lazar, *EMBO Rep* **2005**, *6*, 445.
- [24] C. Chang, J. D. Norris, H. Gron, L. A. Paige, P. T. Hamilton, D. J. Kenan, D. Fowlkes, D. P. McDonnell, *Mol Cell Biol* **1999**, *19*, 8226.
- [25] P. Le Goff, M. M. Montano, D. J. Schodin, B. S. Katzenellenbogen, *J Biol Chem* **1994**, *269*, 4458.
- [26] a)T. Chen, *Curr Opin Chem Biol* **2008**, *12*, 418; b)Y. Shang, M. Brown, *Science* **2002**, *295*, 2465; c)A. K. Shiau, D. Barstad, P. M. Loria, L. Cheng, P. J. Kushner, D. A. Agard, G. L. Greene, *Cell* **1998**, *95*,

- 927; d)A. K. Shiau, D. Barstad, J. T. Radek, M. J. Meyers, K. W. Nettles, B. S. Katzenellenbogen, J. A. Katzenellenbogen, D. A. Agard, G. L. Greene, *Nat Struct Biol* **2002**, *9*, 359; e)P. Germain, J. Iyer, C. Zechel, H. Gronemeyer, *Nature* **2002**, *415*, 187; f)P. Webb, P. Nguyen, P. J. Kushner, *J Biol Chem* **2003**, *278*, 6912; g)H. Greschik, R. Flaig, J. P. Renaud, D. Moras, *J Biol Chem* **2004**, *279*, 33639; h)C. L. Smith, B. W. O'Malley, *Endocr Rev* **2004**, *25*, 45; i)E. K. Keeton, M. Brown, *Mol Endocrinol* **2005**, *19*, 1543; j)V. C. Jordan, B. W. O'Malley, *J Clin Oncol* **2007**, *25*, 5815.
- [27] a)Z. Wang, G. Benoit, J. Liu, S. Prasad, P. Aarnisalo, X. Liu, H. Xu, N. P. Walker, T. Perlmann, *Nature* **2003**, *423*, 555; b)Y. Shi, *Drug Discov Today* **2007**, *12*, 440; c)G. Sun, R. T. Yu, R. M. Evans, Y. Shi, *Proc Natl Acad Sci U S A* **2007**, *104*, 15282.
- [28] a)S. Green, P. Walter, V. Kumar, A. Krust, J. M. Bornert, P. Argos, P. Chambon, *Nature* **1986**, *320*, 134; b)G. L. Greene, P. Gilna, M. Waterfield, A. Baker, Y. Hort, J. Shine, *Science* **1986**, *231*, 1150.
- [29] a)G. G. Kuiper, E. Enmark, M. Peltto-Huikko, S. Nilsson, J. A. Gustafsson, *Proc Natl Acad Sci U S A* **1996**, *93*, 5925; b)S. Mosselman, J. Polman, R. Dijkema, *FEBS Lett* **1996**, *392*, 49.
- [30] J. A. Gustafsson, *J Endocrinol* **1999**, *163*, 379.
- [31] E. H. Kong, A. C. Pike, R. E. Hubbard, *Biochem Soc Trans* **2003**, *31*, 56.
- [32] A. C. Pike, A. M. Brzozowski, R. E. Hubbard, T. Bonn, A. G. Thorsell, O. Engstrom, J. Ljunggren, J. A. Gustafsson, M. Carlquist, *EMBO J* **1999**, *18*, 4608.
- [33] R. K. Dubey, E. K. Jackson, *J Appl Physiol* **2001**, *91*, 1868.
- [34] a)D. P. Edwards, *Annu Rev Physiol* **2005**, *67*, 335; b)R. X. Song, C. J. Barnes, Z. Zhang, Y. Bao, R. Kumar, R. J. Santen, *Proc Natl Acad Sci U S A* **2004**, *101*, 2076; c)J. D. Yager, *J Natl Cancer Inst Monogr* **2000**, *67*.
- [35] A. Haddow, J. M. Watkinson, E. Paterson, *Br. Med. J.* **1944**, *2*, 393.
- [36] K. McKeage, M. P. Curran, G. L. Plosker, *Drugs* **2004**, *64*, 633.
- [37] C. K. Osborne, H. Zhao, S. A. Fuqua, *J Clin Oncol* **2000**, *18*, 3172.
- [38] a)A. U. Buzdar, C. Marcus, F. Holmes, V. Hug, G. Hortobagyi, *Oncology* **1988**, *45*, 344; b)W. Gradishar, J. Glusman, Y. Lu, C. Vogel, F. J. Cohen, G. W. Sledge, Jr., *Cancer* **2000**, *88*, 2047.
- [39] a)B. S. Katzenellenbogen, M. M. Montano, K. Ekena, M. E. Herman, E. M. McInerney, *Breast Cancer Res Treat* **1997**, *44*, 23; b)V. C. Jordan, *Nat Rev Drug Discov* **2003**, *2*, 205.
- [40] *Lancet* **2005**, *365*, 1687.
- [41] a)J. D. Norris, L. A. Paige, D. J. Christensen, C. Y. Chang, M. R. Huacani, D. Fan, P. T. Hamilton, D. M. Fowlkes, D. P. McDonnell, *Science* **1999**, *285*, 744; b)N. Normanno, M. Di Maio, E. De Maio, A. De Luca, A. de Matteis, A. Giordano, F. Perrone, *Endocr Relat Cancer* **2005**, *12*, 721; c)R. Schiff, S. A. Massarweh, J. Shou, L. Bharwani, G. Arpino, M. Rimawi, C. K. Osborne, *Cancer Chemother Pharmacol* **2005**, *56 Suppl 1*, 10; d)R. Schiff, C. K. Osborne, *Breast Cancer Res* **2005**, *7*, 205.
- [42] A. D. Mooradian, J. E. Morley, S. G. Korenman, *Endocr Rev* **1987**, *8*, 1.
- [43] P. M. Matias, P. Donner, R. Coelho, M. Thomaz, C. Peixoto, S. Macedo, N. Otto, S. Joschko, P. Scholz, A. Wegg, S. Basler, M. Schafer, U. Egner, M. A. Carrondo, *J Biol Chem* **2000**, *275*, 26164.
- [44] B. He, L. W. Lee, J. T. Minges, E. M. Wilson, *J Biol Chem* **2002**, *277*, 25631.
- [45] a)B. He, J. T. Minges, L. W. Lee, E. M. Wilson, *J Biol Chem* **2002**, *277*, 10226; b)R. Chmelar, G. Buchanan, E. F. Need, W. Tilley, N. M. Greenberg, *Int J Cancer* **2007**, *120*, 719.
- [46] a)N. Fujimoto, S. Yeh, H. Y. Kang, S. Inui, H. C. Chang, A. Mizokami, C. Chang, *J Biol Chem* **1999**, *274*, 8316; b)H. Y. Kang, S. Yeh, N. Fujimoto, C. Chang, *J Biol Chem* **1999**, *274*, 8570; c)S. Yeh, C. Chang, *Proc Natl Acad Sci U S A* **1996**, *93*, 5517.
- [47] E. Estebanez-Perpina, J. M. Moore, E. Mar, E. Delgado-Rodrigues, P. Nguyen, J. D. Baxter, B. M. Buehrer, P. Webb, R. J. Fletterick, R. K. Guy, *J Biol Chem* **2005**, *280*, 8060.
- [48] E. Langley, J. A. Kempainen, E. M. Wilson, *J Biol Chem* **1998**, *273*, 92.
- [49] M. E. van Royen, S. M. Cunha, M. C. Brink, K. A. Mattern, A. L. Nigg, H. J. Dubbink, P. J. Verschure, J. Trapman, A. B. Houtsmuller, *J Cell Biol* **2007**, *177*, 63.
- [50] a)P. Alen, F. Claessens, G. Verhoeven, W. Rombauts, B. Peeters, *Mol Cell Biol* **1999**, *19*, 6085; b)T. Slagsvold, I. Kraus, T. Bentzen, J. Palvimo, F. Saatcioglu, *Mol Endocrinol* **2000**, *14*, 1603.
- [51] a)C. W. Gregory, B. He, R. T. Johnson, O. H. Ford, J. L. Mohler, F. S. French, E. M. Wilson, *Cancer Res* **2001**, *61*, 4315; b)B. He, R. T. Gampe, Jr., A. T. Hnat, J. L. Faggart, J. T. Minges, F. S. French, E. M. Wilson, *J Biol Chem* **2006**, *281*, 6648.
- [52] B. J. Furr, B. Valcaccia, B. Curry, J. R. Woodburn, G. Chesterson, H. Tucker, *J Endocrinol* **1987**, *113*, R7.
- [53] R. Neri, E. Peets, A. Watnick, *Biochem Soc Trans* **1979**, *7*, 565.
- [54] a)T. Yoshida, H. Kinoshita, T. Segawa, E. Nakamura, T. Inoue, Y. Shimizu, T. Kamoto, O. Ogawa, *Cancer Res* **2005**, *65*, 9611; b)K. Steketeet, L. Timmerman, A. C. Ziel-van der Made, P. Doesburg, A. O. Brinkmann, J. Trapman, *Int J Cancer* **2002**, *100*, 309; c)T. Hara, J. Miyazaki, H. Araki, M. Yamaoka, N. Kanzaki, M. Kusaka, M. Miyamoto, *Cancer Res* **2003**, *63*, 149; d)M. E. Taplin, *Nat Clin Pract Oncol* **2007**, *4*, 236.
- [55] S. M. Dehm, D. J. Tindall, *J Biol Chem* **2006**, *281*, 27882.

- [56] D. M. Lonard, B. W. O'Malley, *Trends Biochem Sci* **2005**, *30*, 126.
- [57] D. M. Lonard, B. W. O'Malley, *Cell* **2006**, *125*, 411.
- [58] L. I. McKay, J. A. Cidlowski, *Mol Endocrinol* **1998**, *12*, 45.
- [59] P. Lefebvre, *Curr Drug Targets Immune Endocr Metabol Disord* **2001**, *1*, 153.
- [60] L. A. Paige, D. J. Christensen, H. Gron, J. D. Norris, E. B. Gottlin, K. M. Padilla, C. Y. Chang, L. M. Ballas, P. T. Hamilton, D. P. McDonnell, D. M. Fowlkes, *Proc Natl Acad Sci U S A* **1999**, *96*, 3999.
- [61] S. Wang, C. Zhang, S. K. Nordeen, D. J. Shapiro, *J Biol Chem* **2007**, *282*, 2765.
- [62] a)D. J. van de Wijngaart, M. E. van Royen, R. Hersmus, A. C. Pike, A. B. Houtsmuller, G. Jenster, J. Trapman, H. J. Dubbink, *J Biol Chem* **2006**, *281*, 19407; b)C. L. Hsu, Y. L. Chen, S. Yeh, H. J. Ting, Y. C. Hu, H. Lin, X. Wang, C. Chang, *J Biol Chem* **2003**, *278*, 23691; c)M. S. Ozers, B. D. Marks, K. Gowda, K. R. Kupcho, K. M. Ervin, T. De Rosier, N. Qadir, H. C. Eliason, S. M. Riddle, M. S. Shekhani, *Biochemistry* **2007**, *46*, 683.
- [63] C. Y. Chang, J. Abdo, T. Hartney, D. P. McDonnell, *Mol Endocrinol* **2005**, *19*, 2478.
- [64] Y. Chen, C. L. Sawyers, H. I. Scher, *Curr Opin Pharmacol* **2008**, *8*, 440.
- [65] N. Heldring, M. Nilsson, B. Buehrer, E. Treuter, J. A. Gustafsson, *Mol Cell Biol* **2004**, *24*, 3445.
- [66] E. H. Kong, N. Heldring, J. A. Gustafsson, E. Treuter, R. E. Hubbard, A. C. Pike, *Proc Natl Acad Sci U S A* **2005**, *102*, 3593.
- [67] a)C. Y. Chang, D. P. McDonnell, *Mol Endocrinol* **2002**, *16*, 647; b)A. R. Bapat, D. E. Frail, *J Steroid Biochem Mol Biol* **2003**, *86*, 143.
- [68] a)R. K. Guy, J. D. Baxter, B. Darimont, W. Feng, R. J. Fletterick, P. J. Kushner, R. L. Wagner, *PCT/US 2001/20969 [WO 2002/02488]*; b)T. R. Geistlinger, R. K. Guy, *J Am Chem Soc* **2003**, *125*, 6852; c)T. R. Geistlinger, A. C. Reynolds, R. K. Guy, *Chem Biol* **2004**, *11*, 273; d)J. P. Northrop, D. Nguyen, S. Piplani, S. E. Olivian, S. T. Kwan, N. F. Go, C. P. Hart, P. J. Schatz, *Mol Endocrinol* **2000**, *14*, 605.
- [69] A. M. Leduc, J. O. Trent, J. L. Wittliff, K. S. Bramlett, S. L. Briggs, N. Y. Chirgadze, Y. Wang, T. P. Burris, A. F. Spatola, *Proc Natl Acad Sci U S A* **2003**, *100*, 11273.
- [70] A. K. Galande, K. S. Bramlett, J. O. Trent, T. P. Burris, J. L. Wittliff, A. F. Spatola, *ChemBiochem* **2005**, *6*, 1991.
- [71] D. Shao, T. J. Berrodin, E. Manas, D. Hauze, R. Powers, A. Bapat, D. Gonder, R. C. Winneker, D. E. Frail, *J Steroid Biochem Mol Biol* **2004**, *88*, 351.
- [72] A. L. LaFrate, J. R. Gunther, K. E. Carlson, J. A. Katzenellenbogen, *Bioorg Med Chem* **2008**, *16*, 10075.
- [73] J. R. Gunther, Y. Du, E. Rhoden, I. Lewis, B. Revennaugh, T. W. Moore, S. H. Kim, R. Dingleline, H. Fu, J. A. Katzenellenbogen, *J Biomol Screen* **2009**, *14*, 181.
- [74] A. L. Rodriguez, A. Tamrazi, M. L. Collins, J. A. Katzenellenbogen, *J Med Chem* **2004**, *47*, 600.
- [75] A. A. Parent, J. R. Gunther, J. A. Katzenellenbogen, *J Med Chem* **2008**, *51*, 6512.
- [76] J. R. Gunther, T. W. Moore, M. L. Collins, J. A. Katzenellenbogen, *ACS Chem Biol* **2008**, *3*, 282.
- [77] J. R. Gunther, A. A. Parent, J. A. Katzenellenbogen, *ACS Chem Biol* **2009**, *4*, 435.
- [78] H. B. Zhou, M. L. Collins, J. R. Gunther, J. S. Comminos, J. A. Katzenellenbogen, *Bioorg Med Chem Lett* **2007**, *17*, 4118.
- [79] a)J. Becerril, A. D. Hamilton, *Angew Chem Int Ed Engl* **2007**, *46*, 4471; b)J. M. Davis, L. K. Tsou, A. D. Hamilton, *Chem Soc Rev* **2007**, *36*, 326.
- [80] A. B. Williams, P. T. Weiser, R. N. Hanson, J. R. Gunther, J. A. Katzenellenbogen, *Org Lett* **2009**, *11*, 5370.
- [81] Y. Wang, N. Y. Chirgadze, S. L. Briggs, S. Khan, E. V. Jensen, T. P. Burris, *Proc Natl Acad Sci U S A* **2006**, *103*, 9908.
- [82] a)L. A. Arnold, A. Kosinski, E. Estebanez-Perpina, R. J. Fletterick, R. K. Guy, *J Med Chem* **2007**, *50*, 5269; b)E. Estebanez-Perpina, L. A. Arnold, N. Jouravel, M. Togashi, J. Blethrow, E. Mar, P. Nguyen, K. J. Phillips, J. D. Baxter, P. Webb, R. K. Guy, R. J. Fletterick, *Mol Endocrinol* **2007**, *21*, 2919.
- [83] R. J. Flower, *Nat Rev Drug Discov* **2003**, *2*, 179.
- [84] L. E. Braverman, R. D. Utiger, *Lippincott Williams & Wilkins, Philadelphia* **2000**.
- [85] E. Estebanez-Perpina, L. A. Arnold, P. Nguyen, E. D. Rodrigues, E. Mar, R. Bateman, P. Pallai, K. M. Shokat, J. D. Baxter, R. K. Guy, P. Webb, R. J. Fletterick, *Proc Natl Acad Sci U S A* **2007**, *104*, 16074.
- [86] H. Faus, B. Haendler, *Biomed Pharmacother* **2006**, *60*, 520.
- [87] a)R. D. Ward, N. L. Weigel, *Biofactors* **2009**, *35*, 528; b)N. L. Weigel, N. L. Moore, *Mol Endocrinol* **2007**, *21*, 2311.
- [88] a)T. Kouzarides, *EMBO J* **2000**, *19*, 1176; b)C. Wang, M. Fu, R. H. Angeletti, L. Siconolfi-Baez, A. T. Reutens, C. Albanese, M. P. Lisanti, B. S. Katzenellenbogen, S. Kato, T. Hopp, S. A. Fuqua, G. N. Lopez, P. J. Kushner, R. G. Pestell, *J Biol Chem* **2001**, *276*, 18375.
- [89] a)K. Xu, H. Shimelis, D. E. Linn, R. Jiang, X. Yang, F. Sun, Z. Guo, H. Chen, W. Li, X. Kong, J. Melamed, S. Fang, Z. Xiao, T. D. Veenstra, Y. Qiu, *Cancer Cell* **2009**, *15*, 270; b)A. Rodriguez-Gonzalez, K. Cyrus, M. Salcius, K. Kim, C. M. Crews, R. J. Deshaies, K. M. Sakamoto, *Oncogene* **2008**, *27*, 7201.

- [90] a)M. Fu, C. Wang, J. Wang, X. Zhang, T. Sakamaki, Y. G. Yeung, C. Chang, T. Hopp, S. A. Fuqua, E. Jaffray, R. T. Hay, J. J. Palvimo, O. A. Janne, R. G. Pestell, *Mol Cell Biol* **2002**, *22*, 3373; b)L. Zhu, N. C. Santos, K. H. Kim, *Endocrinology* **2009**, *150*, 5586; c)J. H. Lee, S. M. Park, O. S. Kim, C. S. Lee, J. H. Woo, S. J. Park, E. H. Joe, I. Jou, *Mol Cell* **2009**, *35*, 806; d)A. R. Daniel, C. A. Lange, *Proc Natl Acad Sci U S A* **2009**, *106*, 14287; e)S. Mukherjee, M. Thomas, N. Dadgar, A. P. Lieberman, J. A. Iniguez-Lluhi, *J Biol Chem* **2009**, *284*, 21296.
- [91] M. Marino, P. Ascenzi, *Steroids* **2008**, *73*, 853.
- [92] M. Sanchez, N. Picard, K. Sauve, A. Tremblay, *Trends Endocrinol Metab* **2009**.
- [93] D. Chen, T. Riedl, E. Washbrook, P. E. Pace, R. C. Coombes, J. M. Egly, S. Ali, *Mol Cell* **2000**, *6*, 127.
- [94] J. Cheng, C. Zhang, D. J. Shapiro, *Endocrinology* **2007**, *148*, 4634.
- [95] H. Lee, W. Bai, *Mol Cell Biol* **2002**, *22*, 5835.
- [96] a)C. Q. Sheeler, D. W. Singleton, S. A. Khan, *Endocr Res* **2003**, *29*, 237; b)Y. Masuhiro, Y. Mezaki, M. Sakari, K. Takeyama, T. Yoshida, K. Inoue, J. Yanagisawa, S. Hanazawa, W. O'Malley B, S. Kato, *Proc Natl Acad Sci U S A* **2005**, *102*, 8126; c)M. Dutertre, C. L. Smith, *Mol Endocrinol* **2003**, *17*, 1296; d)S. Medunjanin, A. Hermani, B. De Servi, J. Grisouard, G. Rincke, D. Mayer, *J Biol Chem* **2005**, *280*, 33006.
- [97] S. F. Arnold, J. D. Obourn, H. Jaffe, A. C. Notides, *J Steroid Biochem Mol Biol* **1995**, *55*, 163.
- [98] D. Chen, P. E. Pace, R. C. Coombes, S. Ali, *Mol Cell Biol* **1999**, *19*, 1002.
- [99] a)E. Castano, D. P. Vorobjekina, A. C. Notides, *Biochem J* **1997**, *326 (Pt 1)*, 149; b)P. B. Joel, J. Smith, T. W. Sturgill, T. L. Fisher, J. Blenis, D. A. Lannigan, *Mol Cell Biol* **1998**, *18*, 1978; c)Y. M. Shah, B. G. Rowan, *Mol Endocrinol* **2005**, *19*, 732; d)K. M. Coleman, M. Dutertre, A. El-Gharbawy, B. G. Rowan, N. L. Weigel, C. L. Smith, *J Biol Chem* **2003**, *278*, 12834; e)H. W. Tsai, J. A. Katzenellenbogen, B. S. Katzenellenbogen, M. A. Shupnik, *Endocrinology* **2004**, *145*, 2730.
- [100] C. C. Valley, R. Metivier, N. M. Solodin, A. M. Fowler, M. T. Mashek, L. Hill, E. T. Alarid, *Mol Cell Biol* **2005**, *25*, 5417.
- [101] G. Bunone, P. A. Briand, R. J. Miksicek, D. Picard, *EMBO J* **1996**, *15*, 2174.
- [102] a)R. A. Campbell, P. Bhat-Nakshatri, N. M. Patel, D. Constantinidou, S. Ali, H. Nakshatri, *J Biol Chem* **2001**, *276*, 9817; b)M. Sun, J. E. Paciga, R. I. Feldman, Z. Yuan, D. Coppola, Y. Y. Lu, S. A. Shelley, S. V. Nicosia, J. Q. Cheng, *Cancer Res* **2001**, *61*, 5985.
- [103] a)A. H. Talukder, D. Q. Li, B. Manavathi, R. Kumar, *Oncogene* **2008**, *27*, 5233; b)W. Zwart, A. Griekspoor, V. Berno, K. Lakeman, K. Jalink, M. Mancini, J. Neeffjes, R. Michalides, *EMBO J* **2007**, *26*, 3534.
- [104] Y. Cui, M. Zhang, R. Pestell, E. M. Curran, W. V. Welshons, S. A. Fuqua, *Cancer Res* **2004**, *64*, 9199.
- [105] a)M. R. Yudt, D. Vorobjekina, L. Zhong, D. F. Skafar, S. Sasson, T. A. Gasiewicz, A. C. Notides, *Biochemistry* **1999**, *38*, 14146; b)M. H. Al-Daheri, B. G. Rowan, *Nucl Recept Signal* **2006**, *4*, e007.
- [106] a)Q. X. Zhang, A. Borg, D. M. Wolf, S. Oesterreich, S. A. Fuqua, *Cancer Res* **1997**, *57*, 1244; b)A. Tremblay, G. B. Tremblay, F. Labrie, V. Giguere, *Mol Cell* **1999**, *3*, 513; c)V. S. Likhite, F. Stossi, K. Kim, B. S. Katzenellenbogen, J. A. Katzenellenbogen, *Mol Endocrinol* **2006**, *20*, 3120.
- [107] a)C. W. Wong, C. McNally, E. Nickbarg, B. S. Komm, B. J. Cheskis, *Proc Natl Acad Sci U S A* **2002**, *99*, 14783; b)F. Barletta, C. W. Wong, C. McNally, B. S. Komm, B. Katzenellenbogen, B. J. Cheskis, *Mol Endocrinol* **2004**, *18*, 1096; c)D. Haas, S. N. White, L. B. Lutz, M. Rasar, S. R. Hammes, *Mol Endocrinol* **2005**, *19*, 2035; d)R. Rajhans, R. K. Vadlamudi, *Clin Exp Metastasis* **2006**, *23*, 1; e)B. J. Cheskis, J. Greger, N. Cooch, C. McNally, S. McLarney, H. S. Lam, S. Rutledge, B. Mekonnen, D. Hauze, S. Nagpal, L. P. Freedman, *Steroids* **2008**, *73*, 901.
- [108] a)L. Varricchio, A. Migliaccio, G. Castoria, H. Yamaguchi, A. de Falco, M. Di Domenico, P. Giovannelli, W. Farrar, E. Appella, F. Auricchio, *Mol Cancer Res* **2007**, *5*, 1213; b)A. Migliaccio, G. Castoria, M. Di Domenico, A. Ciociola, M. Lombardi, A. De Falco, M. Nanayakkara, D. Bottero, R. De Stasio, L. Varricchio, F. Auricchio, *Ann N Y Acad Sci* **2006**, *1089*, 194.
- [109] F. Acconcia, P. Ascenzi, A. Bocedi, E. Spisni, V. Tomasi, A. Trentalance, P. Visca, M. Marino, *Mol Biol Cell* **2005**, *16*, 231.
- [110] A. Tremblay, V. Giguere, *J Steroid Biochem Mol Biol* **2001**, *77*, 19.
- [111] A. Tremblay, G. B. Tremblay, C. Labrie, F. Labrie, V. Giguere, *Endocrinology* **1998**, *139*, 111.
- [112] X. Cheng, R. N. Cole, J. Zaia, G. W. Hart, *Biochemistry* **2000**, *39*, 11609.
- [113] X. Cheng, G. W. Hart, *J Biol Chem* **2001**, *276*, 10570.
- [114] N. Picard, C. Charbonneau, M. Sanchez, A. Licznar, M. Busson, G. Lazennec, A. Tremblay, *Mol Endocrinol* **2008**, *22*, 317.
- [115] K. Sauve, J. Lepage, M. Sanchez, N. Heveker, A. Tremblay, *Cancer Res* **2009**, *69*, 5793.
- [116] A. Migliaccio, G. Castoria, M. Di Domenico, A. de Falco, A. Bilancio, M. Lombardi, M. V. Barone, D. Ametrano, M. S. Zannini, C. Abbondanza, F. Auricchio, *EMBO J* **2000**, *19*, 5406.
- [117] J. B. Biggins, J. T. Koh, *Curr Opin Chem Biol* **2007**, *11*, 99.
- [118] a)J. T. Koh, *Chem Biol* **2002**, *9*, 17; b)J. T. Koh, J. B. Biggins, *Curr Top Med Chem* **2005**, *5*, 413.

- [119] a)R. Lafont, L. Dinan, *J Insect Sci* **2003**, *3*, 7; b)Y. Shi, J. T. Koh, *J Am Chem Soc* **2002**, *124*, 6921; c)D. F. Doyle, D. A. Braasch, L. K. Jackson, H. E. Weiss, M. F. Boehm, D. J. Mangelsdorf, D. R. Corey, *J Am Chem Soc* **2001**, *123*, 11367; d)R. Tedesco, J. A. Thomas, B. S. Katzenellenbogen, J. A. Katzenellenbogen, *Chem Biol* **2001**, *8*, 277; e)A. Q. Hassan, J. T. Koh, *J Am Chem Soc* **2006**, *128*, 8868; f)W. Weber, M. Fussenegger, *J Gene Med* **2006**, *8*, 535.
- [120] a)A. Hashimoto, Y. Shi, K. Drake, J. T. Koh, *Bioorg Med Chem* **2005**, *13*, 3627; b)S. L. Swann, J. Bergh, M. C. Farach-Carson, C. A. Ocasio, J. T. Koh, *J Am Chem Soc* **2002**, *124*, 13795; c)A. Q. Hassan, J. T. Koh, *Angew Chem Int Ed Engl* **2008**, *47*, 7280.
- [121] a)K. Chockalingam, Z. Chen, J. A. Katzenellenbogen, H. Zhao, *Proc Natl Acad Sci U S A* **2005**, *102*, 5691; b)L. J. Schwimmer, P. Rohatgi, B. Azizi, K. L. Seley, D. F. Doyle, *Proc Natl Acad Sci U S A* **2004**, *101*, 14707.
- [122] P. Gallinari, A. Lahm, U. Koch, C. Paolini, M. C. Nardi, G. Roscilli, O. Kinzel, D. Fattori, E. Muraglia, C. Toniatti, R. Cortese, R. De Francesco, G. Ciliberto, *Chem Biol* **2005**, *12*, 883.
- [123] Z. Chen, H. Zhao, *J Mol Biol* **2005**, *348*, 1273.
- [124] a)Y. Shi, J. T. Koh, *Chembiochem* **2004**, *5*, 788; b)K. H. Link, F. G. Cruz, H. F. Ye, E. O'Reilly K, S. Dowdell, J. T. Koh, *Bioorg Med Chem* **2004**, *12*, 5949; c)W. Lin, C. Albanese, R. G. Pestell, D. S. Lawrence, *Chem Biol* **2002**, *9*, 1347.
- [125] K. H. Link, Y. Shi, J. T. Koh, *J Am Chem Soc* **2005**, *127*, 13088.
- [126] a)S. Tanimoto, S. Matsumura, K. Toshima, *Chem Commun (Camb)* **2008**, 3678; b)A. Suzuki, K. Tsumura, T. Tsuzuki, S. Matsumura, K. Toshima, *Chem Commun (Camb)* **2007**, 4260.
- [127] A. R. Schneekloth, M. Pucheault, H. S. Tae, C. M. Crews, *Bioorg Med Chem Lett* **2008**, *18*, 5904.
- [128] M. Awais, M. Sato, X. Lee, Y. Umezawa, *Angew Chem Int Ed Engl* **2006**, *45*, 2707.
- [129] a)R. Paulmurugan, S. S. Gambhir, *Proc Natl Acad Sci U S A* **2006**, *103*, 15883; b)R. Paulmurugan, A. Tamrazi, J. A. Katzenellenbogen, B. S. Katzenellenbogen, S. S. Gambhir, *Mol Endocrinol* **2008**, *22*, 1552.
- [130] A. R. Buskirk, Y. C. Ong, Z. J. Gartner, D. R. Liu, *Proc Natl Acad Sci U S A* **2004**, *101*, 10505.
- [131] M. E. Hahn, T. W. Muir, *Trends Biochem Sci* **2005**, *30*, 26.
- [132] C. M. Yuen, S. J. Rodda, S. A. Vokes, A. P. McMahon, D. R. Liu, *J Am Chem Soc* **2006**, *128*, 8939.
- [133] a)T. W. Muir, D. Sondhi, P. A. Cole, *Proc Natl Acad Sci U S A* **1998**, *95*, 6705; b)R. R. Flavell, T. W. Muir, *Acc Chem Res* **2009**, *42*, 107; c)T. W. Muir, *Biopolymers* **2008**, *90*, 743.

Chapter 2

On-Bead Peptide Screening for Modulators of the Androgen Receptor – Cofactor Interaction

Abstract: Inhibitors of the Nuclear Receptor (NR) – Cofactor interaction represent a novel strategy to modulate NR signaling. The development of an on-bead screening technique allowed the efficient and rapid screening of peptide libraries to identify novel cofactor binding inhibitors (CBIs) for the Androgen Receptor.

2.1 Introduction

Nuclear receptors (NRs) are transcription factors typically under the control of small hydrophobic ligands that bind to the NR ligand binding domain (LBD)^[1]. Depending on the molecular characteristics of these ligands, the LBD^[2] changes its conformation and interacts with cofactor proteins. Corepressors typically bind the antagonist-occupied NR and coactivators will be recruited by the agonist-bound LBD^[3]. The interaction of the LBD with protein coactivators, inducing gene transcription, occurs via leucine-rich motifs in the coactivator proteins. Each of these motifs has the form of a short amphiphatic helix with a characteristic LXXLL^[4] peptide sequence that interacts with almost all nuclear receptors. The Androgen Receptor (AR) represents an exception and preferentially binds to cofactor proteins featuring an FXXLF motif^[5]. This protein-protein interaction constitutes an attractive alternative interface, possibly amendable by synthetic inhibitors that directly disrupt the NR – cofactor interaction (Figure 1). Several synthetic cofactor binding inhibitors (CBIs) of the NR – coactivator interaction have been described in the last years, both based on peptides^[3e, 6] and small molecules^[7]. Peptide based screening methods have relied on large and diverse libraries such as those achieved via phage display techniques^[8] and on smaller focussed libraries obtained via synthetic techniques allowing for the incorporation of non-natural amino acids^[7i, 9]. For the screening of new peptide inhibitor motifs for the NR-Coactivator interaction and also for the generation of peptide binders targeting other NR surfaces, a rapid screening technique providing the opportunity to screen a large library combined with the possibility to incorporate non-natural amino acids would be desirable.

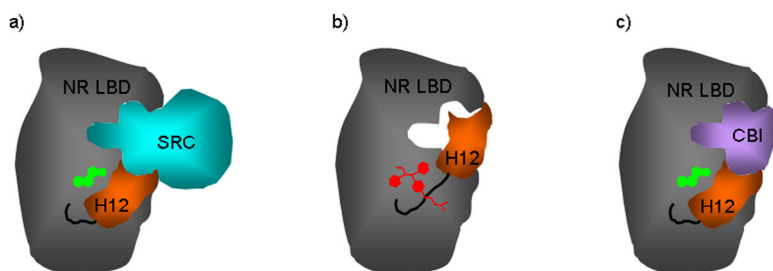


Figure 1: Schematic representation of traditional and CBI antagonism of a nuclear receptor (NR). (a) Conformation of the agonist-bound NR LBD with helix 12 (H12) forming part of the steroid receptor coactivator (SRC) binding site. (b) Conformation of the antagonist-bound NR in which helix 12 occupies the SRC binding site, disrupting the NR – SRC interaction indirectly. (c) Conformation of the agonist-bound NR in which a CBI occupies the SRC binding site, disrupting the NR - SRC interaction directly.

In this chapter the AR^[10] was used as a model system to demonstrate the successful development and evaluation of an One-Bead peptide library screening approach^[11]. Beads amenable to organic synthetic modifications and compatible to protein screening conditions were modified in a combinatorial fashion with a specific peptide library, leading to an One-Bead-One-Compound library (OBOC). Incubation of this peptide library with the fluorescently labelled agonist-bound AR LBD enabled the detection of potential binding events via fluorescence microscopy (Figure 2). It could be shown that sequence analysis of the peptide hits and further evaluation of their binding affinity for the AR LBD provides the entry to identify new peptide binders for NR - coactivator interactions (Figure 2).

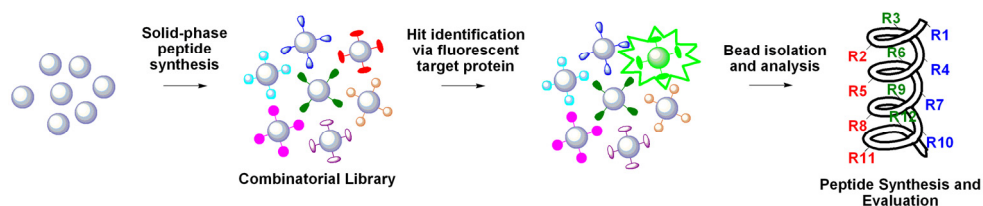


Figure 2: General method for the on-bead screening. Synthesis of an on-bead library of the AR FXXLF motif, incubation with fluorescently labeled AR LBD and detection of the peptide hits under the microscope, followed by isolation and identification of the peptide sequence. The exact binding affinity was analyzed after peptide re-synthesis using a fluorescence polarization competition assay.

2.2 Design and Synthesis of a Combinatorial Peptide Library

In order to evaluate the potential of an OBOC approach to screen for NR binding peptides, it was first focused on a library based on a peptide binding motif for which AR activity was annotated before. The natural sequence of the AR N-terminal domain (SKTYRGAFQNLFQSVREVIG), which binds to the AR LBD via intramolecular interactions, was selected as scaffold to generate a focussed peptide library of potential AR binders^[5a, 12]. Amino acid mutations of the three hydrophobic positions of the characteristic FXXLF motif, the two phenylalanines and the leucine, would result in a library that allows for the evaluation of the hydrophobic groove on the AR with respect to recognition motif and adaptability. The amino acids in these three important positions were exchanged by the six natural amino acids Leu, Phe, Trp, His, Tyr and Met and the four hydrophobic non-natural amino acids (2-thio)phenylalanine (ThioPh), (2-benzothiazolyl)alanine (Bza), (2-naphthyl)alanine (Nap) and (4-chloro)phenylalanine (PhCl) (Figure 3). The hydrophobic nature for the non-natural amino acids was chosen, as

2.3 On-Bead Screening and Biochemical Peptide Hit Validation

In order to screen the combinatorial peptide library on the beads, the AR LBD requires to be modified with a fluorescent tracer. For this, the AR LBD was expressed in *E. coli* as a fusion protein with an N-terminal Glutathione-S-Transferase (GST)^[14] in the presence of the endogenous agonist dihydrotestosterone (DHT) and purified via glutathione affinity chromatography^[10d, 15]. The purity of the fusion protein was confirmed by SDS-PAGE^[16] (Figure 4). Subsequently, the purified AR LBD was conjugated with an amine reactive Texas Red dye and separated from the free dye using size exclusion chromatography (SEC). In order to avoid potential binding interactions promoted by the fluorophore, the average degree of labeling was determined and only preparations were used where the ratio of fluorescent molecules to protein molecules is not higher than 1.0.

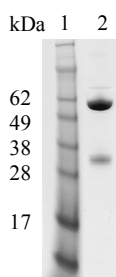


Figure 4: SDS-PAGE gel (15%, Coomassie stained), lane 1: molecular weight marker, lane 2: GST-AR LBD₆₆₄₋₉₁₉ (calcd mass: 51050 Da). The smaller band with a size of around 30 kDa most probably presents a degradation product.

Before screening the bead-displayed peptide library, several control experiments were carried out to evaluate the efficiency of the screening assay. Beads displaying peptides with the unmodified FXXLL motif of the AR N-terminus were used as positive references and beads bearing the LXXLL peptide sequence of SRC 1 Box 2 were used as negative references. Each batch of beads was first blocked by incubation with a large excess of cleared *E. coli*. As reported previously^[17] the use of the diverse mixture of proteins present in an *E. coli* lysate is a particularly effective source of non-specific competitor proteins, which has the effect of competing binding of relatively hydrophobic molecules in the library to the fluorescently-labelled protein. This is desirable, because while these 'sticky' ligands can often exhibit good affinity, they are rarely specific. However, the use of bovine serum albumin (BSA) was found out to be suitable as well. The beads were then incubated with the labelled dihydrotestosterone (DHT)-bound GST-

AR LBD in the presence of a 1000-fold excess of the cleared bacterial lysate or BSA, again to compete non-specific interactions. A 10-fold excess of unlabelled GST is also added to discourage the isolation of peptides that bind this part of the fusion protein. After removing the aqueous solution followed by a washing step, potential AR LBD binding events were visualized under a fluorescence microscope using the optical filter for Texas Red with excitation wavelengths of 555-590 nm and an emission wavelength of 620 nm.

Beads that bind to the labelled AR protein were expected to feature a characteristic bright fluorescence, whereas beads showing no or only weak binding activity for the AR should not light up. After analysis under the fluorescence microscope, the unmodified beads exhibited almost no background fluorescence prior exposure of the protein, as expected (Figure 5a). After modifying the beads with NR specific peptides the intrinsic fluorescence did not significantly increase, indicating the usefulness of these beads for the screening of combinatorial peptide libraries (Figure 5b-c). Next, the beads were analyzed after incubation with the GST-AR LBD-Texas Red. The fluorescence signal for the unmodified beads did not change significantly, indicating the absence of unspecific interactions of the fluorescently GST-AR LBD with the bead. In contrast, the peptide displaying beads showed a remarkable increase in fluorescence (Figure 5d-f). It was noticed that, in line with expectations, the fluorescence of the reference beads with the LXXLL peptide was significantly lower than for those reference beads with the FXXLF motif. The AR prefers coactivator sequences containing FXXLF motifs, however, the AR is also able to recruit several LXXLL motifs, although typically with weaker affinity. In order to determine the ability to distinguish peptide hits from non-hits the two types of control beads with the FXXLF and LXXLL reference sequences were mixed with the unmodified control beads and exposed to the GST-AR LBD-Texas Red as illustrated in Figure 5g-h. The weak binding of the AR LBD for the LXXLL motif could be hardly distinguished from the background fluorescence of the unmodified beads. The beads featuring the reference FXXLF motif featured a very good hit to background contrast under the fluorescence microscope indicating the strong and selective interaction of the Texas Red labelled GST-AR LBD with the peptides immobilized on the beads (Figure 5h).

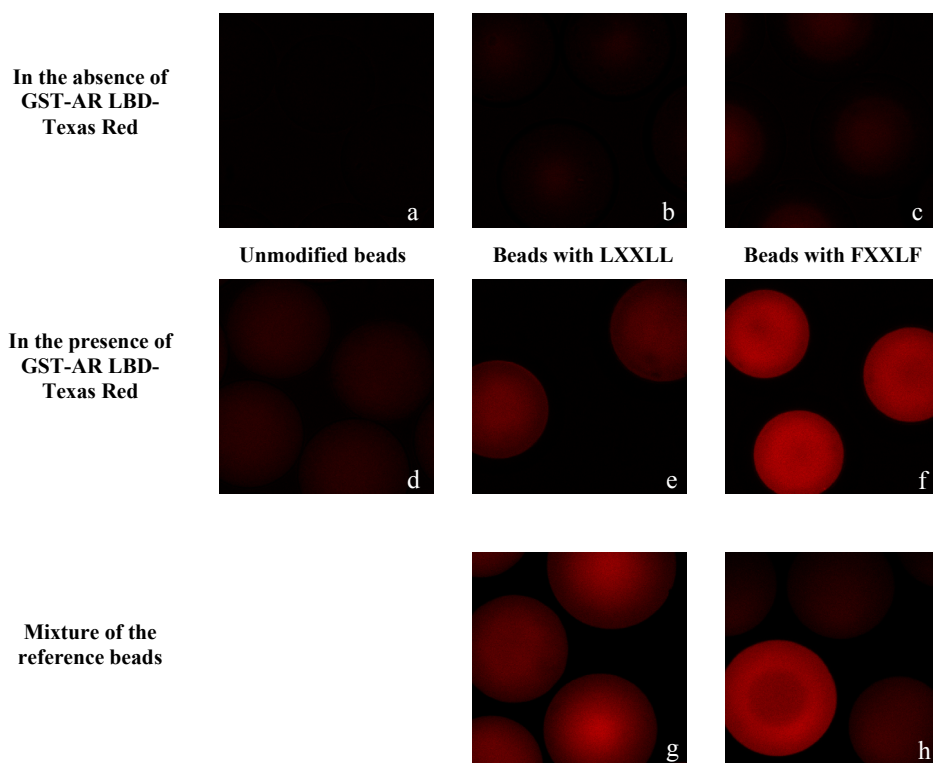


Figure 5: Screening results of the reference beads under different conditions visualized under a fluorescence microscope using a Texas Red filter (λ_{ex} : 555-590 nm; λ_{em} : 620 nm). (a) No background fluorescence was detectable for the unmodified beads. (b) FXXLF peptide displaying beads and (c) LXXLL peptide displaying beads showed no significant intrinsic fluorescence prior to exposure to the protein. (d) Incubation of unmodified beads with the GST-AR LBD-Texas Red resulted in minimal background fluorescence. (e) The negative reference beads displaying the LXXLL containing peptide sequence of SRC 1 box 2 showed a weak binding affinity for the AR LBD resulting in a weak fluorescence signal. (f) Assembling the fluorescently GST-AR LBD on the positive reference beads bearing the natural FXXLL motif for the AR resulted in a high fluorescence of the beads. (g) The mixture of unmodified beads and LXXLL modified beads in the presence of GST-AR LBD-Texas Red (d+e) showed a low, but distinguishable contrast between the hit and non-hit beads. (h) The mixture of unmodified beads and beads with a high affinity for the AR in the presence of the labelled fusion protein (d+f) yielded a very good contrast between the hit and non-hit beads.

The possibility to remove the fluorescently labelled GST-AR LBD again from the bead-bound peptides was investigated using two different methods. One is to simply compete for the peptide – AR LBD interaction with an excess of free peptide. As shown in Figure 6a, the incubation of FXXLF peptide displaying beads bound to the GST-AR LBD Texas Red with a 100-fold excess of free FXXLF peptide for one hour already caused a strong decrease in

fluorescence as a result of AR LBD displacement. The other method to remove the fusion protein from the bead library is the treatment with protein denaturing reagents. Incubating the protein-bound beads with 8 M urea resulted in a significant decrease of the fluorescence signal due to the denaturation of the fusion protein (Figure 6b).

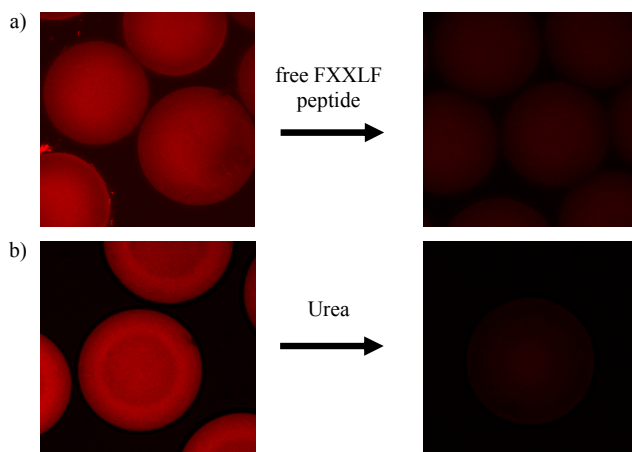


Figure 6: Displacement of the GST-AR LBD Texas Red from the FXXLF peptide displaying beads (a) via competition with free FXXLF peptide and (b) via treatment with the denaturant Urea. The success of removing the fusion protein from the beads was confirmed via the clear decrease of the fluorescence.

The chemical modification of the GST-AR LBD with the Texas Red derivative turned out to lead to a rather unstable protein construct that showed rapid aggregation and denaturation. The fluorescent fusion protein lost its activity after freezing at $-80\text{ }^{\circ}\text{C}$ and remained active only for a few days when stored at $4\text{ }^{\circ}\text{C}$. Therefore, an alternative labelling strategy was developed using the native, not chemically modified AR LBD as target. The already incorporated GST-fusion protein, which confers stability to the AR LBD, was used to specifically label the protein with quantum dots (Qdots) that were functionalized with an α -GST-antibody. The large wavelength difference between the excitation and emission spectrum of the Qdots (λ_{ex} : 405 nm; λ_{em} : 655 nm) has the additional advantage that a potential autofluorescence (high-level and broad-wavelength) of the polystyrene-based resin is not hampering the fluorescence microscopy evaluation of the protein binding to the beads^[17-18]. Further, this labelling strategy prevents the isolation of peptides that recognize only a chemically-modified form of the target protein. Figure 7 illustrates the suitability of the alternative labelling strategy using Qdots to screen on-bead peptide libraries. The signal to noise ratio of the fluorescence signal turned out to be even

better for the α -GST coupled Qdots bound to the GST-AR LBD. Therefore, all later experiments were performed using the Qdots strategy for hit detection.

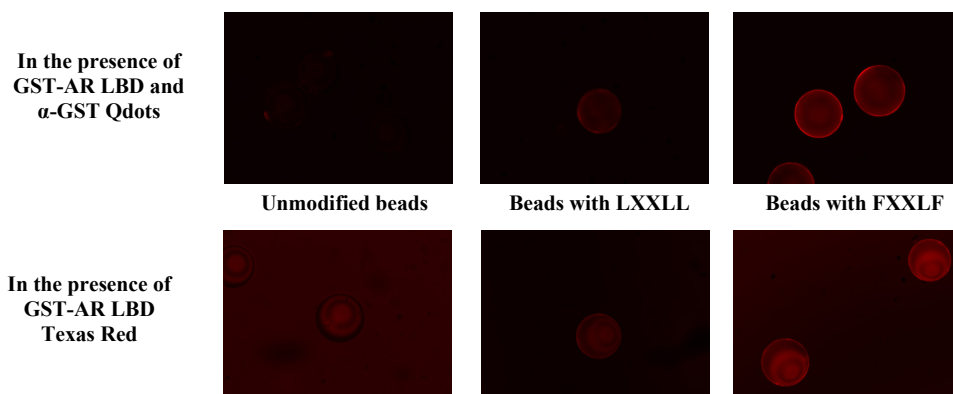


Figure 7: Screening results of reference beads in the presence of Texas Red labelled GST-AR LBD in comparison with α -GST coupled Qdots bound to GST-AR LBD visualized under a fluorescence microscope using a suitable band pass filter (λ_{ex} : 555-590 nm; λ_{em} : 620 nm for Texas Red and λ_{ex} : 405 nm; λ_{em} : 655 nm for the α -GST coupled Qdots).

The bead library with the focused changes in the natural FXXLF motif was screened based on the reasoning that high bead fluorescence intensities are directly correlated with a strong binding affinity of the immobilized peptide sequences for the AR LBD. The brightest beads were isolated from the library for further analysis using a micropipette. Beads were considered as positive hits when the average minimum intensity of the bead, as measured across the bead diagonal, was higher than the intensity of the reference beads with the FXXLF motif. Additional, beads with intensities lower than this reference value (with the intensity around 400) were picked to establish the correlation between bead fluorescence intensity and the affinity of the peptide for the AR LBD. Single beads, visually identified as hits, were incubated with 1% hot SDS to denaturize the bound protein. This was critical to the success of the sequence analysis. The cleavage of the peptide from the single bead was achieved by treatment with 1 M NaOH, which was afterwards neutralized with an equimolar amount of a 1 M HCl solution and formic acid. The sequence of the cleaved peptide was determined using MALDI-TOF mass spectrometry.

Table 1 shows an overview of some of the selected peptide sequences with their corresponding bead fluorescence intensities. Three independent bead screens of the

complete peptide library were performed and around 20 beads per screen were isolated including bad and good hits. The on-bead assay of the focussed library resulted in a diverse set of peptide hits in which the characteristic FXXLF motif was exchanged by various amino acids both of natural and non-natural nature. Surprisingly, most of the identified peptide sequences contained one or two non-natural amino acids at the three randomized amino acid positions. The most abundant amino acid found in position +1 of the FXXLF motif was a leucine (14 x). The most frequently identified amino acid in position +4 was a tyrosine (11x). Interestingly, AR binding motifs featuring a tyrosine in position +4 could be identified in previously studies using phage display^[5b]. The second common amino acid in position +4 was a leucine (9 x), in line with the leucine in the natural FXXLF motif. Interestingly, the prevalent amino acid found in position +5 was (4-chloro)phenylalanine (14 x). This non-natural amino acid is similarly bulky as the natural phenylalanine, indicating the importance of an aromatic bulky residue in this position. The resulting consensus sequence L XX Y PhCl was furthermore the only sequence that was identified more than once as hit in these screening series. The bead with the highest fluorescence featured the sequence PhCl XX Bza F. In contrast, beads with fluorescence lower than that for the FXXLF reference peptide, in general, featured peptide sequences with natural, small or polar amino acids, like methionines and histidines, confirming the importance of large, hydrophobic amino acids, in most cases of aromatic nature, at the +1, +4 and +5 positions.

Table 1: Peptide sequences on selected beads and the observed fluorescence intensity of the bead.

Bead No	~X QN X X~	Fluorescence intensity ^a
1 ^b	~F QN LF~	400
2 ^c	~ClPh QN Bza F~	1000
3	~ClPh QN F Bza~	900
4	~Y QN H ClPh~	850
5	~ClPh QN Y ClPh~	800
6	~L QN Bza Bza~	800
7	~Bza QN W L~	800
8 ^d	~L QN Y ClPh~	750
9	~ThioPh QN Y L~	600
10	~M QN F H~	350
11	~Nap QN L M~	250

^a Average minimum intensity of the beads as measured across the diameter. ^bReference sequence. ^cBest found hit. ^dConsensus sequence

In order to validate the identified hits from the on-bead peptide library screen as *bona fide* peptide inhibitors of the AR – coactivator interaction in solution, they were further analyzed using biochemical studies. This is in particular essential as the AR LBD binds the peptides as homodimer. Especially on a multivalent surface, such as on the beads, weak peptide – AR LBD interactions might result in significant overall affinity.^[17] A competitive fluorescence polarization (FP) displacement assay was therefore developed to assess if the hits from the on-bead screen specifically bind the AR LBD – coactivator surface. Therefore, a large scale re-synthesis of the potential peptides hits was performed via solid-phase peptide synthesis (SPPS) using standard Fmoc-chemistry and HOBt/DIC as coupling reagents on Rink amide resin. Subsequently, the peptides were cleaved under acid conditions and purified by reversed phase preparative high performance liquid chromatography (HPLC). Further, the reference FXXLF peptide was synthesized with an N-terminal cysteine in order to label it with an activated carboxyfluorescein. However, as it turned out that the purified reference peptide was badly soluble in the assay buffer, an alternative AR binding peptide was selected. The peptide NH₂-SSRFESLFAGEKESR-CO₂H obtained by Fletterick *et al.*^[5b] via phage display is one of the best known AR surface binders. After SPPS, this peptide showed a good solubility in the protein buffer. An analogue of the new FXXLF reference peptide featuring an N-terminal cysteine was synthesized and labelled with fluorescein. The binding affinity of the labelled peptide FI-CSSRFESLFAGEKESR was determined in a dose response FP assay. When the small fluorescent peptide is excited with polarized light, it emits fluorescence with polarization inversely proportional to its molecular rotation. If the FI-FXXLF peptide binds to the AR LBD, the molecular rotation of the complex decreases, while the polarization increases. In contrast, titration of the complex of FI-FXXLF peptide and AR LBD with free and unlabelled FXXLF reference peptide results in a concentration dependent competition between free peptide and the tracer and the polarization value decreases accordingly (Figure 8).

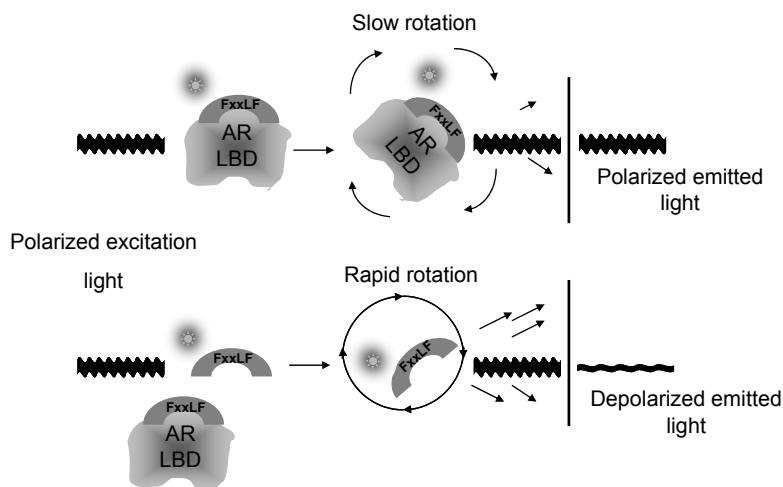


Figure 8: Schematic representation of the fluorescence polarization competition assay. The binding of the fluorescently labeled FXXLF peptide to the androgen receptor results in a large complex, which features slow rotation and high polarization of the emitted light. Displacement of the fluorescent peptide by an inhibitor peptide results in a small fluorescent species with high tumbling rate and subsequent depolarized emission.

The polarization assay was used to determine the dissociation constant K_D of the labelled peptide and the AR LBD. The fluorescent peptide was titrated with increasing concentrations of the GST-AR LBD fusion protein. The K_D values from three independent measurements were obtained through a nonlinear least squares fit to a single-site binding model (Figure 9a). The FXXLF peptide featured a mean K_D of $\sim 0.96 \mu\text{M}$ in excellent agreement with the K_D determined by Fletterick *et al.*^[5b]. To investigate the competitive inhibition of cofactor peptide binding, an AR LBD concentration of $1 \mu\text{M}$ was chosen. At the given concentration of $0.1 \mu\text{M}$ FXXLF peptide and $1 \mu\text{M}$ AR LBD, 50% of the FXXLF peptide exists as complex with the AR LBD. The competitive inhibition assay was carried out by titration of the complex of labelled peptide and DHT-bound AR LBD with the unlabelled FXXLF reference peptide in a concentration-dependent manner (Figure 9b). The affinity of the inhibiting FXXLF peptide for the AR LBD (K_I) could be calculated from its observed IC_{50} value^[19]. After fitting the data^[20], an inhibition constant K_I of $3.6 \mu\text{M}$ was obtained, demonstrating that the chosen competitive FP binding assay conditions are ideal for identifying new small molecule inhibitors of the AR LBD – cofactor interaction.

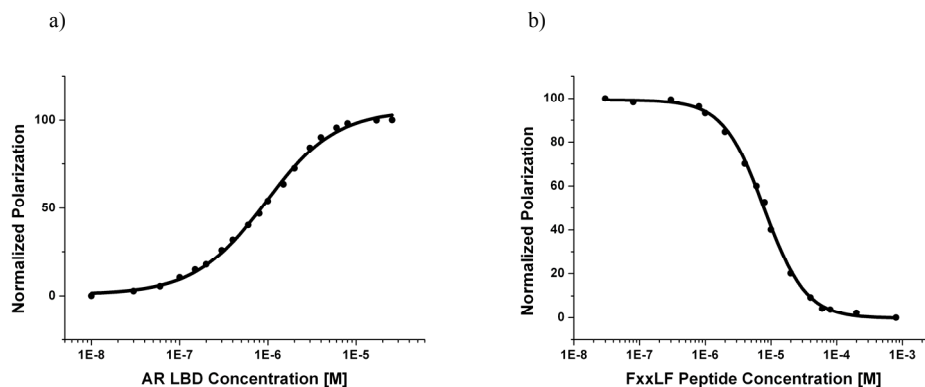


Figure 9: (a) Titration of a constant concentration (0.1 μM) of the fluorescein-labelled FXXLF peptide with increasing concentrations of the ligand saturated AR LBD to determine the K_D ($0.96 \mu\text{M} \pm 0.046$) of the protein-peptide pair. (b) Displacement of the fluorescein-labelled FXXLF peptide from the AR LBD by increasing concentrations of unlabelled FXXLF peptide (IC_{50} : 8 μM , K_I : 3.6 μM).

Using the above described assay, the AR binding potential of the re-synthesized peptide hits was evaluated (Figure 10). The binding affinity of the peptides was measured indirectly via the displacement of the alternative reference FI-FXXLF peptide. The corresponding IC_{50} and K_I values are summarized in Table 2.

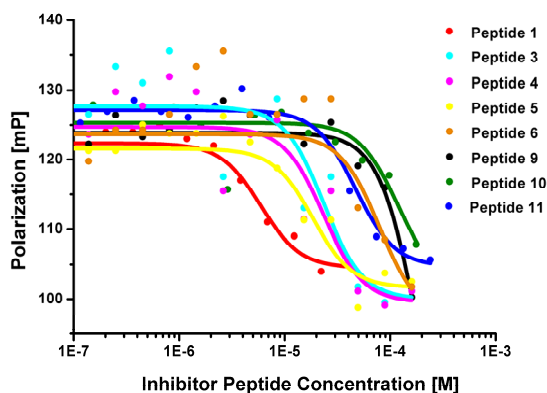


Figure 10: Competitive fluorescence polarization assay of the DHT-bound AR LBD – FI-FXXLF peptide complex (1 μM : 0.1 μM) titrated with increasing concentrations of peptide hits identified as putative AR LBD binders from a combinatorial on-bead library.

Table 2: Sequences and AR LBD binding affinities (IC_{50} and K_I) of a selected set of peptide hits.

Peptide No	~X QN X X~	Fluorescence intensity ^a	IC_{50} [μ M]	K_I [μ M]
1	~F QN LF~ ^b	400	5.3 \pm 1.4	2.24 \pm 0.36
2	~CIPh QN Bza F~ ^c	1000	n.s. ^d	n.s. ^d
3	~CIPh QN F Bza~	900	20 \pm 10	9.35 \pm 4.5
4	~Y QN H CIPh~	850	20 \pm 10	9.35 \pm 4.5
5	~CIPh QN Y CIPh~	800	20 \pm 6.9	9.35 \pm 4.5
6	~L QN Bza Bza~	800	50 \pm 30	23.86 \pm 14.19
7	~Bza QN W L~	800	n.s. ^d	n.s. ^d
8	~L QN Y CIPh~ ^c	725	n.s. ^d	n.s. ^d
9	~ThioPh QN Y L~	600	400 \pm 35	193 \pm 16.6
10	~M QN F H~	350	>890 \pm 108	>430 \pm 51.9
11	~Nap QN L M~	250	40 \pm 9.5	19 \pm 4.28

^a Average minimum intensity of the beads as measured across the diameter.
^bReference sequence AR N-terminus. ^cBest found hit. ^dPeptide was not soluble
^eConsensus sequence

The re-synthesized peptide hits in general featured a poor solubility in the assay buffer, analogous to the FXXLF reference peptide derived from the AR N-terminus. This complicated the determination of the K_I values of the peptide hits. The measured affinities therefore represent only estimations of the real peptide affinities. However, the peptides featured sufficient solubility to displace the FXXLF reference peptide from the AR in the competitive FP inhibition assay. This result indicated that the peptide hits bind the hydrophobic groove on the surface of the AR LBD normally occupied by coactivators. Comparing the K_I values of the measured peptides, a significant difference can be observed between bright beads and beads featuring a low fluorescence. In general, hits that showed a high fluorescence intensity under the microscope, when incubated with the AR and the Qdot featured peptides, also demonstrated higher binding affinities in the competitive inhibition assay. Only peptide 11, featuring the lowest fluorescence, featured a K_I value comparable to the highly fluorescent hits. The lower affinity of the peptide hits in comparison to the reference peptide is in accordance with relatively higher insolubility of these peptides. In order to approve the peptide hits as high affinity binders and thus confirm the efficiency of the peptide library screen, the solubility of the utilized peptides has to be improved. In a recent effort, a similar on-bead peptide library was synthesized based on the sequence of the well-soluble peptide NH_2 -SSRFESLFAAGEKESR-CO₂H obtained by Fletterick *et al.*^[5b] (Göksel *et al.*, unpublished data). Replacement of the AR recognition motif in position +1, +4 and +5 by

natural and non-natural amino acids led to the identification of peptide hits that also featured high binding affinities in the competitive FP inhibition assay. In agreement with the results from the first performed library screen all peptides featured relatively large hydrophobic amino acids at the randomized positions in the peptide, in line with the characteristics of the amino acids at these positions in known interacting motifs^[5b]. Inspection of the amino acids in the randomized positions of the obtained hits reveals that small as well as polar amino acids (A, H, M, P, Q, Y) typically do not occur in the peptide hits. The importance for a large hydrophobic amino acid, in most cases of aromatic nature, at the +1, +4 and +5 positions is thus reconfirmed in this on-bead assay. The best AR binder was obtained with the peptide sequence containing the **NapESChaW** motif (IC₅₀: 0.5 μM) and was even better than the reference peptide (IC₅₀: 0.76 μM) in this study.

2.4 Conclusion

This chapter demonstrates the successful execution of an on-bead library screening for inhibitors of the androgen receptor – coactivator interface. High fluorescently beads were selected via screening the peptide library under the fluorescence microscope and after analysis and re-synthesis of the immobilized peptide sequence the affinity of the potential androgen receptor binders could be determined in a competitive FP inhibition assay. The initial poor solubility of the corresponding peptides could be increased in later experiments by using the well-soluble peptide sequence obtained by Fletterick *et al.*^[5b] as core sequence for the mutagenesis in position +1, +4 and +5 of the FXXLF motif. After the improvement of the solubility of the re-synthesized peptides, the binding affinity of the AR LBD for the peptide on the bead equals the molecular recognition parameters in solution achieving the desired IC₅₀ values. The degree of bead fluorescence thus reflects the binding affinity of the corresponding peptide for the AR LBD.

In general, peptides with a high affinity for the AR LBD featured bulky, hydrophobic amino acids in the randomized positions. Especially large, aromatic, non-natural amino acids turned out to be preferred candidates to allow for AR binding. The usage of non-natural amino acids at the critical interaction positions opens up the opportunity for the increase of membrane permeability of the peptides, the generation of highly AR selective sequences and the design of small molecule CBIs, with similar side-chains functionalities.

All identified hits share the same peptide sequence except the three varied positions, which remarkably could change the affinity to the target protein and indicates the competency to variations using non-natural amino acids in this region. The on-bead assay was thus

established successfully for a focused library in which only three amino acids were exchanged around a well-defined peptide motif.

This screening assay could also be applied to more diverse and randomized peptide sequences and to other NRs than the AR. This will hopefully result in a set of molecular tools that can be used in follow-up studies resulting in the identification of novel inhibitors of NR - cofactor interactions.

2.5 Experimental Section

General Information. Rink Amide MBHA resin with an initial loading of 0.72 mmol/g was purchased from Novabiochem. Fmoc-protected amino acids were purchased from MultiSyntech and Novabiochem in their appropriately protected form. All other reagents were purchased from Aldrich-Sigma, Fluka and Acros. All automated peptide syntheses were performed on a Syro II automated peptide synthesizer (MultiSynTech GmbH), using standard solid phase Fmoc-chemistry. LCMS experiments were performed on an Agilent 1100 series HPLC system connected to a Thermo LCQ Advantage mass spectrometer equipped with an electrospray ion source. Analytical chromatography separations were performed using a C18 Nucleodur gravity column (125 x 4 mm, 3 μ m particle size, Macherey-Nagel). Material was eluted using a gradient system of acetonitrile and water containing 0.1% formic acid and a flow rate of 1 ml/min. Preparative HPLC was performed on a Agilent Series 1100 system equipped with a C18 Nucleodur gravity column (125 x 21 mm, 5 μ m particle size, Macherey-Nagel) using a gradient system of acetonitrile and water each containing 0.1 % trifluoroacetic acid and a flow rate of 25 ml/min. MALDI-TOF mass spectra were recorded on a Voyager DE Pro MALDI-TOF instrument equipped with a LeCroy Digitizer and an internal nitrogen laser using α -cyano-4-hydroxycinnamic acid (CHCA) as Matrix. Initial image acquisition was performed with a confocal fluorescent microscope (Carl Zeiss) using a Texas Red filter (λ_{exc} : 555-590 nm; λ_{em} : 620 nm) or an Axiovert 40 CFL fluorescence microscope (Carl Zeiss). Quantum dots 655 goat anti-GST (Invitrogen) were detected with an excitation wavelength of 405 nm and an emission wavelength of 655 nm and photographed with a camera at an exposure time of 100 ms.

Combinatorial library synthesis. The peptide library synthesis was done by Dr. B. Vaz, Dr. D. Jonkheijm and Dr. H. Göksel. The general scheme for the synthesis of a combinatorial library on beads followed the “split and pool” method as described previously^[21]. For library synthesis 900 mg TentaGel macrobeads with an MBH linker (MB-MBH) and pre-loaded with glycine (loading: 0.237 mmol/g; Rapp Polymer) was split into 10 batches. Double coupling of the next seven amino acids QSVREVI (AA (4 equiv), HBTU (4 equiv), DIPEA (12 equiv) for 40 min) was performed on a Syro II automated peptide synthesizer (MultiSynTech GmbH), using standard solid phase Fmoc-chemistry. After washing the resin twice with DMF one of the ten random amino acids was added to each batch and was manually incorporated in position +5 via double coupling (AA (4 equiv), HBTU (4 equiv), DIPEA (12 equiv) for 40 min), followed by an additional washing step with DMF (4 times). After combining the 10 batches (mixture DCM/DMF) the final Fmoc-deprotection was performed using 40% piperidine/DMF (2.5 mL, 5 min) + 20% piperidine/DMF (2.5 mL, 30 min), followed by a washing step with DMF (6 times). After splitting the resin again into 10 batches the random amino acid in position +4 was incorporated via double coupling (AA (3.6 equiv) + D,L-Ala-NAc (0.4 equiv), HBTU (4 equiv), DIPEA (12 equiv)) for 40 min followed by washing with DMF (2 times). Washing with DMF (4 times) and combination of

the batches was followed by Fmoc-deprotection using 40% piperidine/DMF (2.5 mL, 5 min) + 20% piperidine/DMF (2.5 mL, 30 min), followed by a washing step with DMF (6 times). The next two amino acids QN were again introduced via double coupling (HBTU and D,L-Ala-NAc) using the automated synthesizer. The random amino acid in position 1+ was added manually after batch splitting as described for position 5+ and the amino acid double coupling (HBTU, D,L-Ala-NAc) of S K T Y R G was again carried out using the peptide synthesizer. After the completion of the sequence, the resin was subsequently washed with DMF (5 x 30 s), CH₂Cl₂ (5 x 30 s) and Et₂O (5 x 30 s) and dried under vacuum for 2-3 h. Side chain deprotection was carried out by treatment of the resin for 3 h with a cleavage cocktail containing TFA/TIS/H₂O (95:2.5:2.5)^[22]. Subsequently the resin was washed with DCM (8 x 15 s) and MeOH (8 x 15 s). To prepare the bead library for the screening with the GST-AR LBD the beads were swollen with DMF for 2 h and subsequently incubated with TBST buffer (100 mM Tris/HCl pH 8.0, 150 mM NaCl, 1mM MgCl₂, 0.1% Tween 20) over night at 4 °C.

Peptide Re-synthesis and Purification. All sequences were synthesized from C- to N-terminus on solid support, using an automatic solid phase synthesizer on a 144 μmol scale (200 mg of resin, loading of 0.72 mmol/g). The coupling of amino acids was carried out following standard Fmoc-chemistry, using HOBt/DIC (4 eq.) as amino acid activation, DMF as solvent and 4 eq. of the protected Fmoc-amino acids. The resin was first swollen in DMF (1 x 30 min.) and the Fmoc protecting group was removed by treatment with piperidine/DMF (2/3, 1 x 3 min.; 1/4 1 x 10 min.), then washed with DMF (6 x 1 min.). One cycle of peptide elongation consisted of the following steps. First, the deprotected resin was treated for 50 min with a cocktail containing the appropriate amino acid (4 eq., solution 0.3 M in DMF) with an equimolar addition of HOBt and DIC (4 eq., solution 0.3 M in DMF). After washing the resin with DMF (4 x 1 min.), the Fmoc protecting group was removed by treatment with piperidine/DMF (2/3, 1 x 3 min.; 1/4, 1 x 10 min.). After deprotection, the resin was again washed with DMF (6 x 1 min). These steps were repeated until the desired peptide sequence was complete. After the completion of the sequence, the resin was subsequently washed with DMF (5 x 30 s), CH₂Cl₂ (5 x 30 s) and Et₂O (5 x 30 s) and dried under vacuum for 2-3 h. Cleavage and side chain deprotection was carried out by treatment of the resin for 2 h with a cleavage cocktail containing TFA/H₂O/EDT/TIS (96:2:1:1). The cleaved resin was washed with TFA (2 x 2 ml) and the cleaved peptide was collected, concentrated into less than 1 mL solution and precipitated by addition of cold Et₂O (30 mL). The mixture was cooled, centrifuged (4000 rpm, 5 min, 4 °C) and the Et₂O was decanted from the pellet. Cold Et₂O was added again and the procedure was repeated twice. The crude peptide obtained was dissolved in H₂O/CH₃CN and lyophilized to dryness.

Synthesis and fluorescent labeling of the reference peptide. The peptide SSRFESLFAGEKESR^[5b], and an analogue featuring an N-terminal cysteine reactive towards a fluorescein label (FI-CSSRFESLFAGEKESR) were synthesized on solid support using an automatic solid phase synthesizer (see above). The peptide featuring the N-terminal cysteine was labelled with fluorescein. To a solution of the peptide in potassium phosphate buffer (100 mM, pH 7.2, previously degassed by sonication for 1h at rt) TCEP-HCl (10 eq.) was added and stirred for 1 h at rt. Subsequently, a solution of 4(5)-(Iodoacetamido)fluorescein (5 eq.) in DMSO (10 mg/mL) was added and the homogenous mixture was stirred at rt. The course of the reaction was followed by LCMS. After completion of the reaction, the solution was quenched with an excess of ethanethiol. The resulting mixture was directly lyophilized. The labeled peptide was purified by preparative HPLC. The eluents used were A: H₂O

(+0.1% HCOOH) and B: CH₃CN (+0.1% HCOOH). The method featured a 20 min. gradient of a mixture A/B 90:10 to A/B 50:50 affording the expected fluorescein labeled peptide with high purity (>99 %) and 10% yield for the labeling process. The concentration of the solution of the FI-peptide in HEPES buffer was determined using the UV absorption at 495 nm (molar absorption of fluorescein 75000 M⁻¹ cm⁻¹) by four different dilutions of the original FI-peptide solution, measuring the absorption value at 495 nm (the zero values were corrected by the corresponding values measured at 580 nm).

GST-AR LBD Expression and Purification. The hAR LBD (residues 664-919), subcloned into the vector pGEX-KG with an N-terminal GST tag, was provided by P. Donner (Bayer-Schering Pharma AG). The plasmid was transformed into high-density *Escherichia coli* BL21 (DE3) cells and grown in 6 L TB medium using selection with ampicillin (100 µg/ml). The cultures were incubated at 37 °C to an OD₆₀₀ of ~1.2 and after cooling down to 18 °C protein expression was induced by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 100 µM. The cells were grown in the presence of 10 µM DHT (5α-Androstan-17β-ol-3-one, Fluka) for 18-20 h at 18 °C before harvested by centrifugation (Beckman Coulter, Avanti J26 XP) at 4500 rpm, 20 min. Subsequently, the cells were lysed with a microfluidizer (4 passes at 600 kPa) in buffer H (50 mM HEPES, pH 7.3, 300 mM NaCl, 5 mM EDTA, 10% glycerol, 100 µM DHT, 100 µM PMSF, 10 mM DTT) and centrifuged at 20.000 rpm for 30 min. The soluble cell lysate was immobilized on a glutathione Sepharose 4 Fast Flow affinity matrix (Amersham Biosciences), washed with buffer A (50 mM HEPES, pH 7.3, 300 mM NaCl, 5 mM EDTA, 10% glycerol, 10 µM DHT, 1 mM DTT) and eluted with buffer A containing 15 mM glutathione. Fractions containing the fusion protein were combined and desalted on a Sephadex G25 PD-10 column (Amersham Biosciences) pre-equilibrated with buffer A. Purity and characterization of the eluted fractions was established by SDS-PAGE using a molecular weight marker (PageRuler PlusPrestained Protein Ladder, Fermentas) and photometric determination of protein concentration using Nanodrop at a wavelength of 280 nm. In case of cleavage of the GST moiety with thrombin over night at 4 °C a final purification step was followed using a HiTrap SP cation exchange column (Amersham Biosciences).

Texas Red conjugation to the GST-AR LBD. The DHT-bound GST-AR LBD fusion protein (2 mg/ml) was buffer exchanged to 50 mM HEPES, pH 7.3, 1 mM DTT and 10 µM DHT using a pre-equilibrated Sephadex G25 PD-10 column (Amersham Biosciences) and subsequently adjusted to pH 8.3 with 1 M sodium bicarbonate in a total volume of 550 µl. Texas Red succinimidyl ester (1 mg) was dissolved in 10 µl DMSO and added to the protein solution. After incubation for 1 hour at rt on a rotating wheel, the reaction was quenched by addition of 300 µl hydroxylamine (1.5 M). Subsequently, the dye conjugated protein was separated from the free dye using a Superdex 75 column (Amersham Biosciences; 10 x 300 mm), pre-equilibrated with 1 x PBS (100 ml) and monitoring the separation using a UV-lamp. The labelled protein was eluted from the column using 1 x PBS and the absorbance of the protein (0.3 mg/ml) was measured using a UV/VIS-spectrometer at λ_{250-700 nm} (Cary 100 Bio; Varian). The degree of labeling (DOL) was estimated using the following formula^[17]:

$$DOL = \frac{A_{\max} \times MW}{[protein] \times \epsilon_{dye}} \quad (\text{Eq. 1})$$

where A_{max} is the absorbance of the protein dye conjugate at λ_{max} for the dye (595 nm), MW is the molecular weight of the protein, ϵ_{dye} is the extinction coefficient of the dye at its absorbance maximum (80000) and $[protein]$ is the protein concentration in mg/ml.

On-bead Screening Methodology. The on-bead assay was performed using following procedure. The beads were swollen in DMF for 2 h at room temperature and subsequently rinsed with buffer containing 50 mM HEPES, pH 7.3, 300 mM NaCl, 5 mM EDTA, 10 % glycerol and 10 μ M DHT three times for five minutes, and soaked overnight in the same buffer at 4 °C. Subsequently, the beads were blocked with a 1000-fold excess of *E. coli* lysate or 75 μ M BSA overnight at 4 °C. The beads were then washed thrice with buffer (containing 50 mM HEPES, pH 7.3, 300 mM NaCl, 5 mM EDTA, 10 % glycerol and 10 μ M DHT) and incubated with 0.5 μ M GST-AR LBD fusion protein for 2 h at rt and a 100-fold excess of unlabelled GST. The beads were thoroughly (5 x) washed with buffer. Beads incubated with Texas Red labelled GST-AR LBD were directly visualized under a confocal fluorescence microscope (Carl Zeiss) fitted with a Texas Red filter (λ_{ex} : 555-590 nm; λ_{em} : 620 nm). Beads that were treated with chemically unmodified GST-AR LBD were incubated with 0.01 μ M Qdot goat anti-glutathione-S-transferase (Invitrogen) for 4 h at rt followed by a washing step with buffer (3 x 5 min). The beads were subsequently observed and characterized using a fluorescence microscope (Axiovert 40 CFL fluorescence microscope; Carl Zeiss) with a 405 nm excitation and a 655 nm emission filter and an exposure time of 100 ms. The individual picked beads containing bound protein were then treated with 1 % of a SDS-solution at 99 °C for 30 min and washed to remove any attached protein. After transferring the beads into a new vial, the cleavage of each single bead from the resin was performed using 4.4 μ L of 1 M NaOH. After shaking the solution for 15 min at rt, this solution was neutralized by 4.4 μ L of 1 M HCl and 4.4 μ L formic acid^[17]. The resulting solution was then analyzed via MALDI-TOF using α -cyano-4-hydroxycinnamic acid (CHCA) as matrix.

Fluorescence Polarization Assay hAR LBD. Purified hAR LBD was serially diluted into reaction buffer (20 mM HEPES, pH 7.2, 50 mM KCl, 1 mM EDTA, 10 μ M DHT, 1 mM DTT, 1 mg/ml BSA) with final concentrations of 10 μ M DHT and 0.1 μ M fluorescein-labeled peptide FI-CSSRFESLFLFAGEKESR^[5b] using a black 384-well plate (Perkin Elmer, Optiplate-384 F). The final volume in each well was 50 μ L. Each plate also included the following controls: the empty well, 0.1 μ M fluorescent peptide in reaction buffer and reaction buffer containing the same final concentration of DMSO as used to solve the ligand. The samples were mixed by pipetting up and down and subsequently the plate was centrifuged (5 min, 7000 rpm, 4 °C) to remove bubbles. The reaction was incubated at 4 °C for 1 h in the dark. The fluorescence polarization was measured at 23 °C using a plate reader (Safire², Tecan) with an excitation at 470 nm, an emission at 519 nm, a gain of 80 and 50 reads per well. The G-factor (1.154) was obtained by collecting parallel and perpendicular components of fluorescence from an 8 nM solution of fluorescein in water. The data derived from the fluorescence polarization (in millipolarization, mP) were normalized, plotted against increasing concentrations of the hAR LBD and then fitted with ORIGIN 7 (Scientific Graphing and Analysis Software, OriginLab Corp.) using nonlinear regression analysis with a sigmoidal dose-response (variable slope) equation to determine the K_D value of the fluorescein labeled peptide – hAR LBD complex. With these experimental data and the use of equation 2 a K_D of 0.96 μ M was determined for the fluorescent reference peptide – AR LBD interaction.

$$P = P_0 + \frac{(P_{fin} - P_0)(A_0 + B_0 + K_d - \sqrt{(A_0 + B_0 + K_d)^2 - 4A_0B_0})}{2A_0} \quad (\text{Eq. 2})$$

where P is the measured polarization value, P_0 is the polarization of the free fluorescent ligand, P_{fin} is the polarization of the bound ligand, A_0 is the total concentration of fluorescent peptide, B_0 is the protein concentration and K_D is the dissociation constant of the protein-peptide complex. Each data point represents the average of an experimental condition performed in at least triplicate. Z' factor analysis was performed as recently described by Zhang *et al*^[23]. Assays with a Z' factor between 0.5 and 1.0 are considered to be reliable and robust.

Fluorescence Polarization Competitive Displacement Assay hAR LBD. The optimized assay mixture contained 0.1 μM fluorescein-labeled peptide FI-CSSRFESLFAGEKESR^[5b] and 1 μM of purified hAR LBD in reaction buffer (20 mM HEPES, pH 7.2, 50 mM KCl, 1 mM EDTA, 10 μM DHT, 1 mM DTT, 1mg/ml BSA). Inhibition experiments were performed in 384-well plates (Perkin Elmer, Optiplate-384 F) by adding 40 μl of the reaction mixture to 10 μl of increasing amounts of peptide inhibitor (diluted in reaction buffer). Samples without inhibitor as well as samples containing only FI-labeled peptide in reaction buffer were used as controls. The reaction was incubated at room temperature for 1 h in the dark (measurements repeated after 24 h and stored at 4 $^\circ\text{C}$ showed no significant difference in the IC_{50} values). The fluorescence polarization was measured at 23 $^\circ\text{C}$ using a plate reader (Safire², Tecan) with an excitation at 470 nm and an emission at 519 nm. The polarization data (in millipolarization, mP) were generally normalized, plotted against the \log_{10} of increasing concentrations of the inhibitor and then fitted with a Klotz binding model^[20] to a sigmoid curve using ORIGIN 7.5 (Scientific Graphing and Analysis Software, OriginLab Corp.) to determine the K_I and IC_{50} value of the inhibitor. The competitive binding of the peptides was measured at least two times in independent duplo experiments. The fluorescence polarization data were used to calculate the IC_{50} using equation 3.

$$P = P_{\min} + \frac{(P_{\max} - P_{\min})}{\left(1 + 10^{(\log(x) - \log(\text{IC}_{50}))}\right)} \quad (\text{Eq. 3})$$

where P is the polarization, P_{\min} is the minimum value of polarization and P_{\max} is the maximum value of polarization. With the aid of equation 4, the K_D of the fluorescein labeled peptide and the IC_{50} values of each miniprotein, the K_I values for the different peptides could be determined^[19].

$$K_i = \frac{\text{IC}_{50}}{1 + \frac{A_0 \cdot (y_0 + 2)}{2 \cdot K_d \cdot (y_0 + 1)} + y_0} - K_d \cdot \left(\frac{y_0}{y_0 + 2}\right) \quad (\text{Eq. 4})$$

where IC_{50} is the half maximal inhibitor concentration, K_D is the dissociation constant of the protein - labelled peptide complex and y_0 is the initial bound-to-free concentration ratio for labelled peptide.

$$y_0 = \frac{AB}{A} = \frac{B}{K_D}$$

where AB is the concentration of protein-peptide complex and A the concentration of unbound peptide.

2.6 References

- [1] a)M. J. Tsai, B. W. O'Malley, *Annu Rev Biochem* **1994**, *63*, 451; b)D. J. Mangelsdorf, C. Thummel, M. Beato, P. Herrlich, G. Schutz, K. Umesono, B. Blumberg, P. Kastner, M. Mark, P. Chambon, R. M. Evans, *Cell* **1995**, *83*, 835; c)H. Gronemeyer, J. A. Gustafsson, V. Laudet, *Nat Rev Drug Discov* **2004**, *3*, 950; d)P. Germain, B. Staels, C. Dacquet, M. Spedding, V. Laudet, *Pharmacol Rev* **2006**, *58*, 685; e)J. T. Moore, J. L. Collins, K. H. Pearce, *ChemMedChem* **2006**, *1*, 504.
- [2] R. V. Weatherman, R. J. Fletterick, T. S. Scanlan, *Annu Rev Biochem* **1999**, *68*, 559.
- [3] a)W. B. Pratt, D. O. Toft, *Endocr Rev* **1997**, *18*, 306; b)C. K. Glass, D. W. Rose, M. G. Rosenfeld, *Curr Opin Cell Biol* **1997**, *9*, 222; c)N. J. McKenna, R. B. Lanz, B. W. O'Malley, *Endocr Rev* **1999**, *20*, 321; d)F. Schaufele, X. Carbonell, M. Guerbado, S. Borngraerber, M. S. Chapman, A. A. Ma, J. N. Miner, M. I. Diamond, *Proc Natl Acad Sci U S A* **2005**, *102*, 9802; e)N. B. Mettu, T. B. Stanley, M. A. Dwyer, M. S. Jansen, J. E. Allen, J. M. Hall, D. P. McDonnell, *Mol Endocrinol* **2007**, *21*, 2361.
- [4] a)D. M. Heery, E. Kalkhoven, S. Hoare, M. G. Parker, *Nature* **1997**, *387*, 733; b)R. S. Savkur, T. P. Burris, *J Pept Res* **2004**, *63*, 207.
- [5] a)B. He, J. A. Kempainen, E. M. Wilson, *J Biol Chem* **2000**, *275*, 22986; b)E. Hur, S. J. Pfaff, E. S. Payne, H. Gron, B. M. Buehrer, R. J. Fletterick, *Plos Biology* **2004**, *2*, E274.
- [6] a)T. R. Geistlinger, R. K. Guy, *J Am Chem Soc* **2003**, *125*, 6852; b)A. M. Leduc, J. O. Trent, J. L. Wittliff, K. S. Bramlett, S. L. Briggs, N. Y. Chirgadze, Y. Wang, T. P. Burris, A. F. Spatola, *Proc Natl Acad Sci U S A* **2003**, *100*, 11273; c)T. R. Geistlinger, A. C. McReynolds, R. K. Guy, *Chem Biol* **2004**, *11*, 273; d)A. K. Galande, K. S. Bramlett, J. O. Trent, T. P. Burris, J. L. Wittliff, A. F. Spatola, *Chembiochem* **2005**, *6*, 1991; e)B. Vaz, S. Mocklinghoff, S. Folkertsma, S. Lusher, J. de Vlieg, L. Brunsveld, *Chem Commun (Camb)* **2009**, 5377.
- [7] a)D. Shao, T. J. Berrodin, E. Manas, D. Hauze, R. Powers, A. Bapat, D. Gonder, R. C. Winneker, D. E. Frail, *J Steroid Biochem Mol Biol* **2004**, *88*, 351; b)L. A. Arnold, E. Estebanez-Perpina, M. Togashi, N. Jouravel, A. Shelat, A. C. McReynolds, E. Mar, P. Nguyen, J. D. Baxter, R. J. Fletterick, P. Webb, R. K. Guy, *J Biol Chem* **2005**, *280*, 43048; c)L. A. Arnold, E. Estebanez-Perpina, M. Togashi, A. Shelat, C. A. Ocasio, A. C. McReynolds, P. Nguyen, J. D. Baxter, R. J. Fletterick, P. Webb, R. K. Guy, *Sci STKE* **2006**, *2006*, pl3; d)A. L. Rodriguez, A. Tamrazi, M. L. Collins, J. A. Katzenellenbogen, *J Med Chem* **2004**, *47*, 600; e)I. G. Schullman, R. A. Heyman, *Chem Biol* **2004**, *11*, 639; f)S. Gaillard, M. A. Dwyer, D. P. McDonnell, *Mol Endocrinol* **2007**, *21*, 62; g)L. Wang, W. J. Zuercher, T. G. Consler, M. H. Lambert, A. B. Miller, L. A. Orband-Miller, D. D. McKee, T. M. Willson, R. T. Nolte, *J Biol Chem* **2006**, *281*, 37773; h)J. Becerril, A. D. Hamilton, *Angew Chem Int Ed Engl* **2007**, *46*, 4471; i)E. Estebanez-Perpina, L. A. Arnold, P. Nguyen, E. D. Rodrigues, E. Mar, R. Bateman, P. Pallai, K. M. Shokat, J. D. Baxter, R. K. Guy, P. Webb, R. J. Fletterick, *Proc Natl Acad Sci U S A* **2007**, *104*, 16074; j)J. R. Gunther, T. W. Moore, M. L. Collins, J. A. Katzenellenbogen, *ACS Chem Biol* **2008**, *3*, 282; k)J. R. Gunther, A. A. Parent, J. A. Katzenellenbogen, *ACS Chem Biol* **2009**, *4*, 435; l)A. L. LaFrata, J. R. Gunther, K. E. Carlson, J. A. Katzenellenbogen, *Bioorg Med Chem* **2008**, *16*, 10075; m)A. A. Parent, J. R. Gunther, J. A. Katzenellenbogen, *J Med Chem* **2008**, *51*, 6512; n)A. B. Williams, P. T. Weiser, R. N. Hanson, J. R. Gunther, J. A. Katzenellenbogen, *Org Lett* **2009**, *11*, 5370; o)T. W. Moore, C. G. Mayne, J. A. Katzenellenbogen, *Mol Endocrinol* **2009**; p)J. Y. Hwang, L. A. Arnold, F. Zhu, A. Kosinski, T. J. Mangano, V. Setola, B. L. Roth, R. K. Guy, *J Med Chem* **2009**, *52*, 3892.
- [8] a)C. Chang, J. D. Norris, H. Gron, L. A. Paige, P. T. Hamilton, D. J. Kenan, D. Fowlkes, D. P. McDonnell, *Mol Cell Biol* **1999**, *19*, 8226; b)C. Y. Chang, J. Abdo, T. Hartney, D. P. McDonnell, *Mol Endocrinol* **2005**, *19*, 2478; c)J. M. Hall, C. Y. Chang, D. P. McDonnell, *Mol Endocrinol* **2000**, *14*, 2010.
- [9] a)J. A. Patch, A. E. Barron, *Curr Opin Chem Biol* **2002**, *6*, 872; b)K. H. Pearce, M. A. Iannone, C. A. Simmons, J. G. Gray, *Drug Discov Today* **2004**, *9*, 741; c)C. Denicourt, S. F. Dowdy, *Science* **2004**, *305*, 1411; d)J. R. Gunther, Y. Du, E. Rhoden, I. Lewis, B. Revennaugh, T. W. Moore, S. H. Kim, R. Dingledine, H. Fu, J. A. Katzenellenbogen, *J Biomol Screen* **2009**, *14*, 181.
- [10] a)H. E. MacLean, G. L. Warne, J. D. Zajac, *J Steroid Biochem Mol Biol* **1997**, *62*, 233; b)E. P. Gelmann, *J Clin Oncol* **2002**, *20*, 3001; c)L. Wang, C. L. Hsu, C. Chang, *Prostate* **2005**, *63*, 117; d)W. Gao, C. E. Bohl, J. T. Dalton, *Chem Rev* **2005**, *105*, 3352.

- [11] a)K. S. Lam, S. E. Salmon, E. M. Hersh, V. J. Hraby, W. M. Kazmierski, R. J. Knapp, *Nature* **1991**, 354, 82; b)R. A. Houghten, C. Pinilla, S. E. Blondelle, J. R. Appel, C. T. Dooley, J. H. Cuervo, *Nature* **1991**, 354, 84; c)M. Lebl, V. Krchnak, N. F. Sepetov, B. Seligmann, P. Strop, S. Felder, K. S. Lam, *Biopolymers* **1995**, 37, 177; d)K. S. Lam, D. Lake, S. E. Salmon, J. Smith, M. L. Chen, S. Wade, F. Abdul-Latif, R. J. Knapp, Z. Leblova, R. D. Ferguson, V. V. Krchnak, N. F. Sepetov, M. Lebl, *Methods* **1996**, 9, 482; e)K. S. Lam, M. Lebl, V. Krchnak, *Chem Rev* **1997**, 97, 411; f)K. S. Lam, R. Liu, S. Miyamoto, A. L. Lehman, J. M. Tuscano, *Acc Chem Res* **2003**, 36, 370.
- [12] a)H. J. Dubbink, R. Hersmus, C. S. Verma, H. A. van der Korput, C. A. Berrevoets, J. van Tol, A. C. Ziel-van der Made, A. O. Brinkmann, A. C. Pike, J. Trapman, *Mol Endocrinol* **2004**, 18, 2132; b)K. Stekete, C. A. Berrevoets, H. J. Dubbink, P. Doesburg, R. Hersmus, A. O. Brinkmann, J. Trapman, *European Journal of Biochemistry* **2002**, 269, 5780.
- [13] R. Betney, I. J. McEwan, *J Mol Endocrinol* **2003**, 31, 427.
- [14] R. N. Armstrong, *Chem Res Toxicol* **1991**, 4, 131.
- [15] a)P. M. Matias, P. Donner, R. Coelho, M. Thomaz, C. Peixoto, S. Macedo, N. Otto, S. Joschko, P. Scholz, A. Wegg, S. Basler, M. Schafer, U. Egner, M. A. Carrondo, *J Biol Chem* **2000**, 275, 26164; b)E. R. LaVallie, J. M. McCoy, *Curr Opin Biotechnol* **1995**, 6, 501; c)D. B. Smith, K. S. Johnson, *Gene* **1988**, 67, 31.
- [16] U. K. Laemmli, L. A. Amos, A. Klug, *Cell* **1976**, 7, 191.
- [17] T. Kodadek, K. Bachhawat-Sikder, *Mol Biosyst* **2006**, 2, 25.
- [18] a)K. Bachhawat-Sikder, T. Kodadek, *J Am Chem Soc* **2003**, 125, 9550; b)H. J. Olivos, K. Bachhawat-Sikder, T. Kodadek, *Chembiochem* **2003**, 4, 1242; c)M. Han, X. Gao, J. Z. Su, S. Nie, *Nat Biotechnol* **2001**, 19, 631.
- [19] T. J. Burke, K. R. Loniello, J. A. Beebe, K. M. Ervin, *Comb Chem High Throughput Screen* **2003**, 6, 183.
- [20] I. M. Klotz, J. M. Longfellow, O. H. Johnson, *Science* **1946**, 104, 264.
- [21] P. G. Alluri, M. M. Reddy, K. Bachhawat-Sikder, H. J. Olivos, T. Kodadek, *J Am Chem Soc* **2003**, 125, 13995.
- [22] a)I. Coin, P. Schmieder, M. Bienert, M. Beyermann, *Biopolymers* **2007**, 88, 565; b)V. Cavallaro, P. E. Thompson, M. T. W. Hearn, *Journal of Peptide Science* **2001**, 7, 529.
- [23] J. H. Zhang, T. D. Y. Chung, K. R. Oldenburg, *Journal of Biomolecular Screening* **1999**, 4, 67.

Chapter 3

Miniproteins as Modulators of the Androgen Receptor – Cofactor Interaction

Part of this work has been published: B. Vaz, S. Möcklinghoff, S. Folkertsma, S. Lusher, J. de Vlieg, L. Brunsveld, *Chem. Commun.* **2009**, 36, 5377

Abstract: On the search of novel modulators of the Nuclear Receptor – Cofactor Interaction natural helical miniproteins could be identified as an ideal framework to generate helical coactivator mimics. Introduction of an FXXLF motif into the helix of the miniproteins led to potent and selective androgen receptor binders with a well-defined fold and controlled helix length.

3.1 Introduction

Nuclear receptors (NRs) are multidomain transcription factors that regulate developmental and physiological processes^[1, 2]. Typically, ligands bind in the binding pocket of the NR ligand binding domain (LBD) and modulate the interaction of the NR with a large set of interacting proteins, the so-called coactivators^[3-5]. These proteins bind the agonist-bound state of most NRs with peptide sequences featuring an LXXLL motif^[6] or an FXXLF motif^[7] in case of the androgen receptor (AR). This protein - protein interaction has recently emerged as a possible modulatory interface. Coactivator binding inhibitors (CBIs)^[8] specifically antagonize the NR – coactivator interaction and especially the estrogen (ER) and the thyroid receptor (TR) have been targeted *via* both peptide based^[9-15] and small molecule^[16-25] approaches in this respect.

The AR, as an entry to target for example prostate cancer, has however attracted less attention. Peptide based AR cofactor mimics have been generated *via* phage display^[26] and small molecules have been identified *via* screening^[27, 28]. Design based approaches for peptidic cofactor mimics for the AR have not been reported thus far. A structural analysis of the AR – cofactor interaction shows that the FXXLF motif binds as a short α -helix and is located between a charge clamp at a fixed position on the receptor surface^[29, 30] (Figure 1). The length and stability of the helix thus appear to be crucial elements to achieve optimal binding. Control over helix length and stability is typically difficult for regular peptides and therefore stabilization of short peptides in a defined and stable fold is being pursued^[31-34].

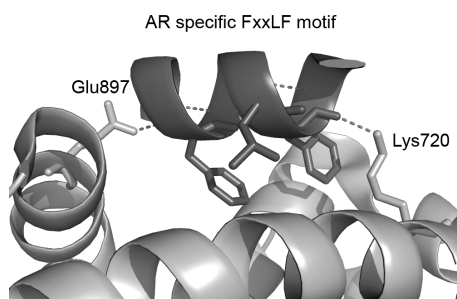


Figure 1: Zoom in on the co-crystal structure of the AR ligand binding domain (LBD) in complex with the AR specific FXXLF motif (PDB 1t7r; black)^[30]. The conserved AR residues E₈₉₇ and K₇₂₀ form the so called ‘charge clamp’ and interact with the coactivator peptide via hydrogen bonds.

Naturally occurring small proteins, so-called miniproteins, possess a well-defined secondary structure and biological activity and are usually stabilized by multiple disulfide

bridges. Examples of natural miniproteins are animal toxins (from bee, scorpion and snake) that affect the ion channel functions of cells. These small structures are conformationally very stable and also compatible with rather different sequences and biological activities. Therefore, miniproteins are particularly attractive structural scaffolds to design protein modulators by grafting the functional residues onto this stable polypeptide scaffold^[35-39]. Due to their small size they can be obtained by chemical synthesis, which allows any desired chemical modification.

In order to gain additional knowledge on the nuclear receptor (NR) – cofactor interaction, the use of miniproteins featuring an α -helix was explored here. Due to the stabilization given by the cysteine bridges, it is possible to introduce mutations conserving the natural helical secondary structure of the original miniprotein. Therefore, it was proposed to incorporate the FXXLF/LXXLL motifs of NR-cofactors in these systems to generate nuclear receptor binders (Figure 2). Here, for the first time a design and synthesis approach is applied to generate peptide binders for the AR with a well-defined fold and controlled helix length. It is demonstrated that natural helical miniprotein scaffolds with differing helical segment lengths provide an ideal framework for the generation of helical AR coactivator mimics, by mutating only three to four amino acids (Figure 2).

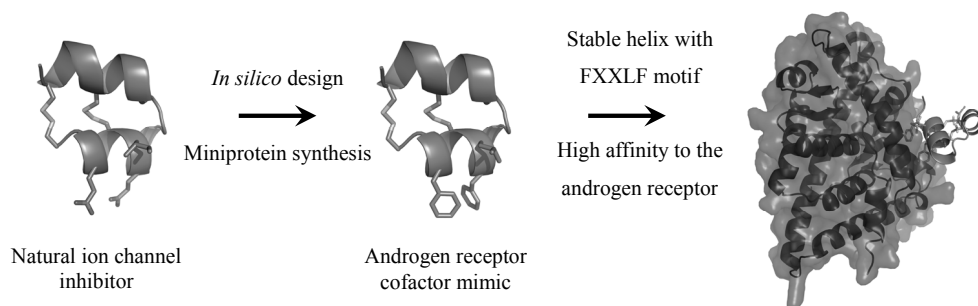


Figure 2: Strategy for the generation of miniproteins binding the androgen receptor (AR) coactivator interaction site, by insertion of the FXXLF motif in the helix of the miniprotein.

3.2 Computational Design and Synthesis of the Miniproteins

In order to identify suitable miniprotein scaffolds featuring stable helices of different lengths, the PDB^[40] was browsed. Several cysteine-stabilized toxins that affect ion channel functions were chosen. The 18-residue neurotoxin Apamin^[41] from the honey bee (*Apis mellifera*) consists of a C-terminal two-turn α -helix cross-linked with two disulfide bonds to the N-terminus. The scorpion toxin κ -hefutoxin1^[42] from *Heterometrus fulvipes* features two anti-parallel α -helices stabilized by two disulfide bridges (PDB: 1hp9). CD 4M3, a scyllatoxin

analogue^[43] derived from *Leiurus quinquestriatus hebraeus*, consists of a double-stranded anti-parallel β -sheet linked with three disulfide bridges to a short α -helix (PDB: 1d5q). The miniprotein Om-toxin 3^[44] from the scorpion *Opisthacanthus madagascariensis* presents a helix-loop-helix fold stabilized by two disulfide bridges (PDB: 1wqd) (Figure 3).

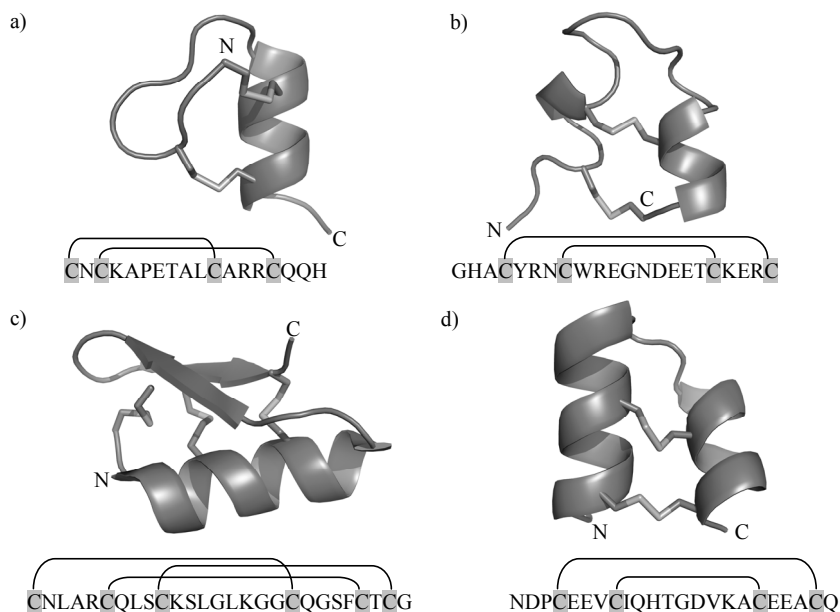


Figure 3: Representation of four different types of natural miniproteins selected as potential scaffolds to design AR modulators. Apamin (a), κ -hefutoxin (b), CD 4M3 (c) and Om-toxin 3 (d) feature stabilized helices by means of two or three disulfide bridges.

In silico, the AR specific FXXLF motif (PDB: 1t7r)^[30] was introduced into the miniprotein helix at positions that did not interfere with the conformational crucial disulfide bridges (Figure 4). The mutated miniproteins were overlaid with the AR ligand binding domain (LBD) surface in complex with a known cofactor motif^[30] to identify a suitable fit, especially with respect to the positioning of the FXXLF motif and to observe the overall fit of the other parts of the miniprotein into the helix binding groove (Figure 4). For certain miniproteins (Het-5, Het-6, Omt-2), the N-terminus of the peptide sequence was found to clash with the surface of the AR. To enable a better fit, these clashing amino acids were deleted from the miniprotein sequence.

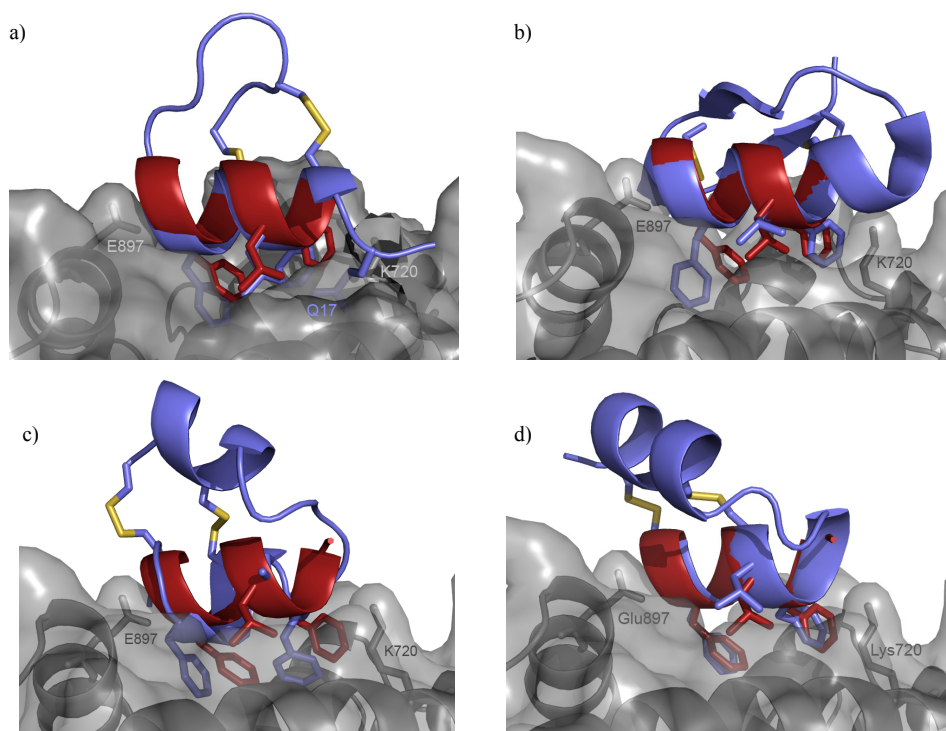


Figure 4: Overlay of chosen mutated miniproteins (blue) with the crystal structure of a linear FXXLF peptide^[30] (red) and the AR LBD surface (grey). (a) The helical segment of the apamin mutant Apa-1 exhibits a good overlay and the three hydrophobic FXXLF side-chains are similarly positioned. Only Gln₁₇ in the C-terminus clashes with the AR surface. (b) The length of the helix in the scyllatoxin mimic Scy-2 could cause significant problems of space in the Lys₇₂₀ region. (c) Deletion of the first two N-terminal amino acids from the natural sequence of hefutoxin results in Het-5 and avoids collision of the miniprotein with the charge clamp residue Glu₈₉₇ of the AR. (d) The Om-toxin mutant Omt-2 features a very good overlay with the linear FXXLF peptide and the AR LBD surface. The deletion of the first three amino acids avoids a clash with the AR surface and enables the shorter helix to be positioned between the charge clamp of Lys₇₂₀ and Glu₈₉₇.

Apart from the FXXLF motif, charge - charge interactions between the cofactor and the surface of the LBD are known to play an important role in peptide binding. For three miniproteins (Het-3, Het-4, Omt-1, Omt-2) it was observed that specific amino acids of the native miniprotein might result in unfavorable charge – charge interactions. For these miniproteins, a specific point mutation, *e.g.* E to R, was inserted (Table 1). Further, the C-terminal bump of the apamin (Apa-1) was observed to be close to the charge clamp amino acid Lys₇₂₀ (Figure 4a). Therefore, it was decided to exchange the Gln₁₇, touching the AR surface, by a glycine to avoid potential clashes. In a few cases, the FXXLF motif was

replaced by an LXXLL motif to investigate potential cross-activity between the AR and other NRs and to validate the need for an FXXLF motif in this miniprotein for AR binding affinity.

Energy minimization was used to analyze, whether minor clashes of the miniprotein with the receptor surface could be overcome by protein flexibility. Additionally, the energy minimization was used to analyze whether a specific point mutation was sterically feasible. This *in silico* exercise resulted in the design of 17 potential AR binding miniproteins (Table 1).

Table 1: Miniproteins, given names and corresponding sequences

Natural Miniprotein	Name	Sequence ^a
Apamin	Apa-0	CNCKAPETALCARRCQQH
	Apa-1	CNCKAPETA F CAL F CQQH
	Apa-2	CNCKAPETA F CAL F CQGH
	Apa-3	CNCKAPETALC ALL CQQH
CD4M3 mimic (scyllatoxin)	Scy-0	CNLARCQLSCKSLGLKGGCQGSFCTCG
	Scy-1	CNL F R C L F SCGSLGLKGGCQGSFCTCG
	Scy-2	CNL A F C Q L F CKSLGLKGGCQGSFCTCG
	Scy-3	CNL A L C Q L L C CKSLGLKGGCQGSFCTCG
κ-hefutoxin	Het-0	GHACYRNCWREGNDEETCKERC
	Het-1	GHACY F N C L F EGNDEETCKERC
	Het-2	GHACYR F C W L F GNDEETCKERC
	Het-3	GHACYRNCWREGNDR F T C L F R C
	Het-4	GHACYRNCWREGNDEF F C K L F C
	Het-5	ACY F N C L F EGNDEETCKERC
	Het-6	ACYR F C W L F GNDEETCKERC
Om-toxin3	Omt-0	NDPCEEVCIQHTGDVKACEEACQ
	Omt-1	N R F C E L F CIQHTGDVKACEEACQ
	Omt-2	C R F V C L F HTGDVKACEEACQ
	Omt-3	NDPCEEVCIQHTGDV F A C L F A C Q
	Omt-4	NDPCEEVCIQHTGDV F C E L F CQ
	Omt-5	NDPCEEVCIQHTGDV L A C L L A C Q

^aDisulfide bridged cysteines are highlighted in grey, inserted mutations are underlined and the signature amino acids are represented in bold.

The 17 FXXLF mutated and *in silico* defined miniprotein sequences and the corresponding four natural unmodified analogues as references were synthesized *via* solid phase peptide synthesis^[45] using a rink amide resin. Peptide elongation involved the treatment

of the resin with piperidine to cleave the Fmoc-protecting group^[46], followed by coupling of the next amino acid with N,N'-diisopropylcarbodiimide (DIC) and hydroxybenzotriazole (HOBt)^[47] (Figure 5). After releasing the linear peptide from the resin^[48] and removing all side-chain protections^[49], the crude product was dissolved in a mixture of trifluoroethanol (TFE) and phosphate buffer. TFE is known to support the formation of α -helices in peptides^[50]. The formation of the stabilizing disulfide bridges was induced by addition of Tris [2-carboxyethyl] phosphine hydrochloride (TCEP) to the peptide solution, to break any previously formed disulfide bridges, and subsequent oxidization by exposure to air for 24 - 48 h^[51]. The complete oxidation of all peptides was confirmed by LC-MS. After removal of the TFE and salts, the oxidized miniproteins were purified by reverse-phase HPLC. Subsequently, the identities and purities of the folded miniproteins were assessed by LC-MS (ESI mass spectrometry). All miniproteins presented the expected molecular mass and purity between 85% and 99%.

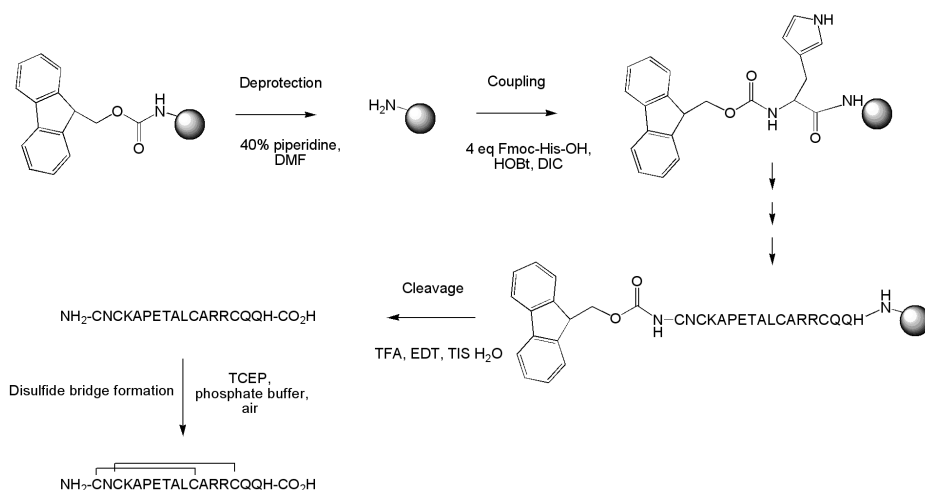


Figure 5: Modular synthesis of miniproteins (example shown for Apa-0) using repeated amino acid coupling and Fmoc deprotection steps followed by resin cleavage and disulfide bridge formation.

To evaluate the structural impact of introducing an FXXLF motif into the miniprotein, circular dichroism (CD)^[52-55] experiments were conducted to compare the α -helicity of the mutated and natural miniproteins. CD measurements in the far UV region provide information about the peptide bond asymmetric environment and reflect the secondary structure content of the designed miniproteins. As expected, the CD spectra of all miniproteins featured the typical shape of peptides and proteins with a high α -helical content characterized by an

absorption maxima at 192 nm and two absorption minima at 208 and 222 nm^[52] (Figure 6). Furthermore, the mutated miniproteins exhibited a far-UV CD spectrum, similar to that presented by the native miniproteins indicating the correct fold of the generated miniproteins. The observed decrease in the CD signal of distinct miniproteins might be caused by impurities and consequential lower miniprotein concentrations. Also the higher content of phenylalanines in the modified miniproteins could account for specific changes in the CD spectrum.

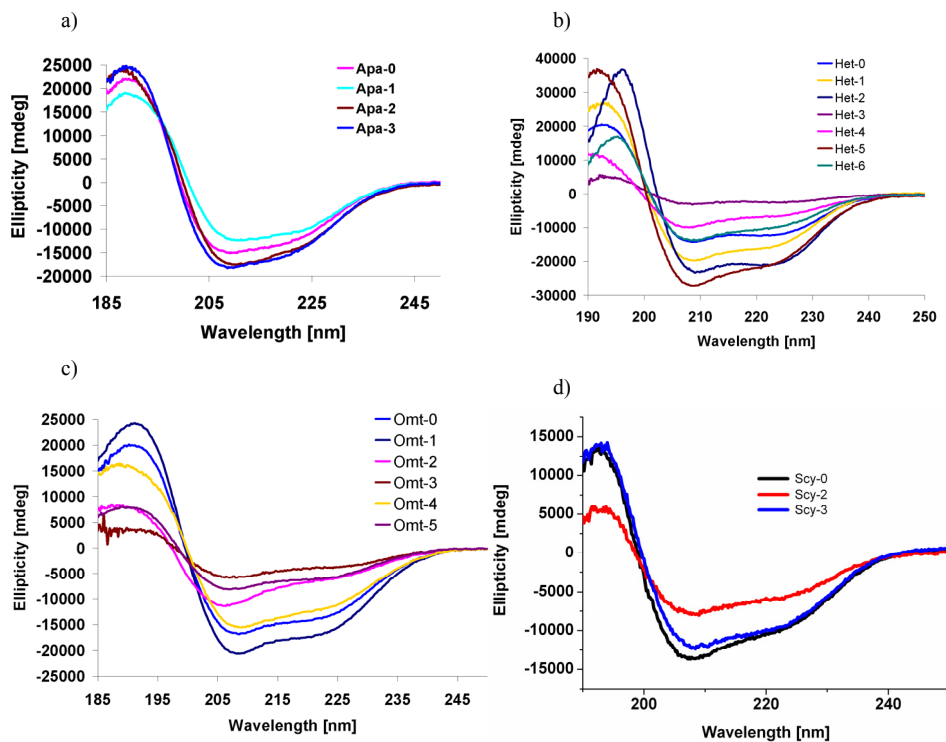


Figure 6: Far UV circular dichroism spectra of 50 μM of the engineered miniproteins recorded in 5 mM potassium buffer pH 7.4^[56]. a) Apamin, b) κ -hefutoxin, c) Om-toxin3, d) CD4M3.

3.3 Androgen Receptor Binding Studies of the Miniproteins

To evaluate the binding affinity of the newly designed miniproteins for the AR LBD, a fluorescence polarization (FP) assay was set up^[57], comparable to protocols previously published for other NRs^[9-13, 16-20]. In this competitive assay, a complex of AR LBD and a fluorescein-labeled FXXLF reference peptide is preformed, and subsequently titrated with increasing concentrations of the miniprotein, possibly displacing the fluorescent peptide^[57].

First, the AR LBD (aa 664 – 919) was expressed in *E. coli* as fusion protein with an N-terminal His₆-tag. However, after purification using ion-metal affinity chromatography

(IMAC), the protein displayed a strong decrease in stability and rapidly aggregated. Therefore, it was decided to express the AR LBD as a fusion protein with an N-terminal glutathione-S-transferase (GST)^[58]. GST is known to increase the conformational stability of the fusion protein^[59]. After purification using glutathione affinity chromatography^[59, 60] and subsequent thrombin cleavage of the GST moiety, cation exchange chromatography (CEX) was performed to remove the cleaved GST and the thrombin from the protein solution. The eluted AR LBD fractions were analyzed by SDS-PAGE^[61] (Figure 7). Expression and all purification steps were performed in the presence of DHT (Dihydrotestosterone), a precursor of the natural AR ligand testosterone, to enable the correct fold of the AR LBD. Later experiments were performed with the GST-AR LBD fusion protein, as the GST had no negative effect on the determination of peptide binding affinities, but even increased the fluorescence polarization signal by increasing the molecular mass of the peptide – protein complex.

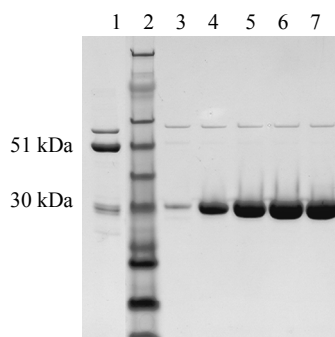


Figure 7: SDS-PAGE gel (15%, Coomassie stained), lane 1: GST-AR LBD before thrombin cleavage (calcd mass: 51050 Da), lane 2: molecular weight marker, lane 3-7: Eluted AR LBD fractions after CEX (calcd. mass: 29962 Da).

The FXXLF sequence of the N-terminal AF-1 of the AR LBD was chosen as AR binding reference peptide. After solid phase peptide synthesis and purification, the peptide sequence NH₂-SKTYRGAFQNLFQSVREVI-CO₂H turned out to be badly soluble. Therefore, it was decided to synthesize a FXXLF peptide obtained by Fletterick *et al.*^[30] via phage display. This peptide with the sequence NH₂-SSRFESLFAGEKESR-CO₂H was well-soluble in buffer and previously shown to bind the AR with high affinity^[30]. This peptide was selected for the competitive fluorescence polarization assay as the competitive control cofactor binding inhibitor (CBI) for the AR LBD. An analogue of this peptide, featuring an N-terminal cysteine, was synthesized, allowing the introduction of a fluorescein label. Therefore, a peptide solution, containing TCEP, was incubated with 4(5)-(iodoacetamido)-

fluorescein. The course of the reaction was followed by LC-MS till no unlabeled peptide could be detected before quenching with ethanethiol. Purification using HPLC afforded the expected fluorescein labeled peptide with high purity (>99 %) and 10% yield for the labeling process.

The affinity of the FXXLF peptide for the ligand occupied AR LBD was measured via FP. FP is a powerful and sensitive technology for the determination of small molecules binding to larger ones in solutions. It monitors changes in the apparent size of fluorescently labelled molecules^[62-64] by measuring the rate of its rotation. Small fluorescent molecules rotate quickly and give a low FP value. Large molecules, on the other hand, rotate more slowly and therefore have higher FP values. When the small fluorescein labelled FXXLF reference peptide is excited with plane-polarized light, the emitted light will be largely depolarized. The peptide rotates rapidly in solution during the fluorescence lifetime (the time between excitation and emission) resulting in a low polarization signal. Interaction of the peptide with the recombinant AR LBD forms a big complex that rotates more slowly than the tracer and increases the FP value. Depending upon the concentration of free FXXLF peptide in the AR LBD solution, a competition between free peptide and the tracer occurs and the polarization value changes accordingly (Figure 8).

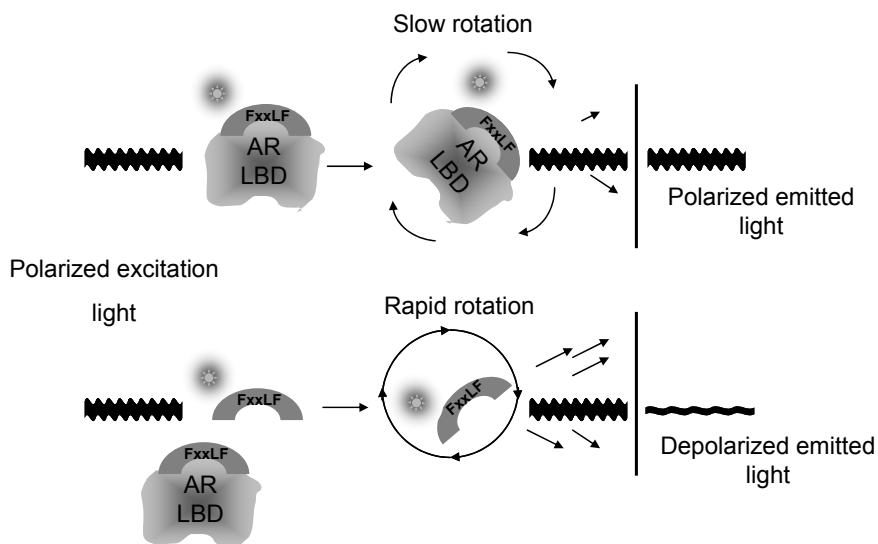


Figure 8: Schematic representation of the fluorescence polarization competition assay. After excitation with polarized light fluorescently labeled reference peptide rotates rapidly and the emitted light is largely depolarized. The binding of the fluorescently labeled FXXLF peptide to the receptor results in a much larger complex. Due to its slower rotation, this complex emits light which remains polarized.

According to Fletterick *et al.*^[30], the apparent dissociation constant K_D for the binding of the selected peptide to the AR LBD was determined at 1.1 μM . In order to determine the K_D of this peptide in our assay, the FP of the labelled peptide was measured as a function of total AR LBD concentration (Figure 9a). As it is generally optimal to use the fluorophore-labelled peptide at a concentration that is at least 10 times below the K_D ^[62], a peptide concentration of 0.1 μM was used. During the measurement it has to be assured that the detected total fluorescence intensity of the fluorophore does not change after association of the peptide tracer with the AR LBD. Differences in the fluorescence intensities of bound and free labelled peptide could lead to potentially significant misinterpretations of the FP data^[65].

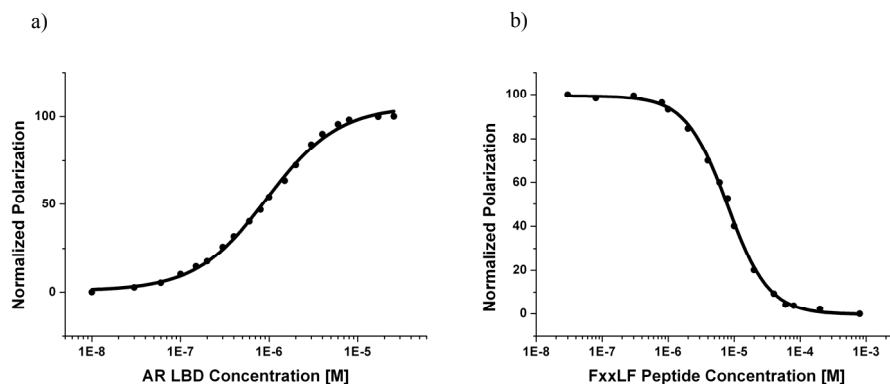


Figure 9: (a) Titration of a constant concentration (0.1 μM) of the fluorescein-labelled FXXLF peptide with increasing concentrations of the ligand saturated AR LBD to estimate the K_D ($0.96 \mu\text{M} \pm 0.046$) of the protein-peptide pair. (b) Displacement of the fluorescein-labelled FXXLF peptide from the AR LBD by increasing concentrations of unlabelled FXXLF peptide (IC_{50} : 8 μM , K_I : 3.6 μM).

The dynamic range (ΔmP) of 213 ± 1.5 and the K_D values of three independent measurements were obtained through a nonlinear least squares fit to a single-site binding model. The FXXLF peptide featured a mean K_D of $\sim 0.96 \mu\text{M}$ in excellent agreement with the K_D determined by Fletterick *et al.*^[30]. FP could thus successfully be established to study the affinity of a fluorescein-labelled FXXLF peptide for the ligand occupied AR LBD. The precision and suitability of the FP assay is also defined by the Z' factor^[66]. The Z' factor is a measure of the statistical effect size in the assay and should exhibit a value close to the maximum value of 1^[66]. The Z' factor for the cofactor peptide binding experiment in Figure 9a was calculated from three independent measurements and was determined at 0.95, confirming that these assay conditions are highly suitable for equilibrium binding experiments.

To investigate the competitive inhibition of cofactor binding, the concentration ratio between the receptor and the K_D of the labelled peptide should be at least 1. Therefore, an AR LBD concentration of 1 μM was chosen. At the given concentration of 0.1 μM FXXLF peptide and 1 μM AR LBD 50% of the FXXLF peptide exists as complex with the AR LBD. A competitive inhibition assay was carried out by titrating the pre-formed complex of labelled peptide and AR LBD with the unlabelled FXXLF peptide in a concentration-dependent manner (Figure 9b). After fitting the data^[67], a dynamic range of 100 ± 1.4 , an inhibition constant K_I of 3.6 μM and a Z' factor of 0.75 was obtained, demonstrating that the chosen competitive FP binding assay conditions are ideal for identifying small molecule inhibitors of the AR LBD – cofactor interaction. The affinity of the inhibiting FXXLF peptide for the AR LBD (K_I) has been calculated from its observed IC_{50} value^[68]. The IC_{50} is the concentration of inhibitor required to displace 50% of the labelled FXXLF peptide.

In order to investigate the AR binding potential and the cofactor binding capacities of the synthesized miniproteins the FP competition assay was performed in triplet with increasing concentrations of each miniprotein solution (Figure 10). The binding affinity of the miniproteins was measured indirectly via the FI-FXXLF peptide displacement. The calculated IC_{50} and K_I values obtained are summarized in the Table 2.

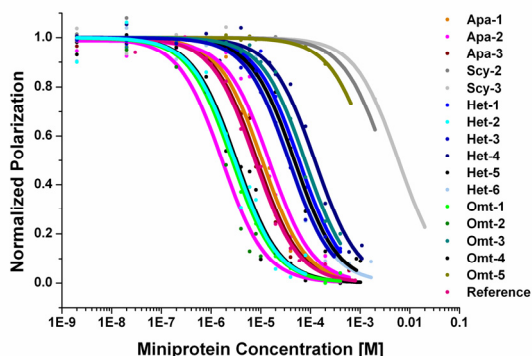


Figure 10: Competitive displacement analysis of the AR LBD – FI-FXXLF peptide complex (1 μM : 0.1 μM) by increasing concentrations of the miniproteins using FP.

Table 2: Miniproteins, sequences and AR binding affinities

Name	Sequence	IC ₅₀ (μM)	K _i (μM)
Reference	SSR FESLF AGEKESR (*)	9 ± 1	3.9 ± 0.5
Apa_0	CNCKAPETAL CARR QQH	n.a.	n.a.
Apa_1	CNCKAPETA FCA LF QQH	10 ± 4	4.4 ± 1.8
Apa_2	CNCKAPETA FCA LF QQH	10 ± 3	4.4 ± 1.8
Apa_3	CNCKAPETAL CALL QQH	9 ± 3	3.9 ± 1.4
Scy_0	CN LARC QLS CK SLGLKGG CQGSFCTCG	n.a.	n.a.
Scy_1	CN LFR LF SGSLGLKGG CQGSFCTCG	(a)	(a)
Scy_2	CN LAF CQLF CKSLGLKGG CQGSFCTCG	≈ 350	≈ 163
Scy_3	CN LALC QLL CKSLGLKGG CQGSFCTCG	≈ 640	≈ 298
Het_0	GHAC YRNC WREGNDEET CKERC	n.a.	n.a.
Het_1	GHAC YFNC LF EGNDEET CKERC	50 ± 10	23 ± 4.6
Het_2	GHAC YR FCW LF EGNDEET CKERC	3 ± 1	1.1 ± 0.4
Het_3	GHAC YRNC WREGNDR FT CLFRC	40 ± 4	18.4 ± 1.8
Het_4	GHAC YRNC WREGNDR RF CKLF C	130 ± 30	60.3 ± 14
Het_5	AC YFNC LF EGNDEET CKERC	5.5 ± 1	2.3 ± 0.5
Het_6	AC YR FCW LF EGNDEET CKERC	45 ± 8	20.7 ± 3.7
Omt_0	NDP CE EV CI QHTGDVKA CEEACQ	n.a.	n.a.
Omt_1	NR FC ELF CI QGTGDVKA CEEACQ	2,6 ± 0,3	0,9 ± 0,1
Omt_2	CR FV CLF H HT GDVKA CEEACQ	1,7 ± 0,3	0,48 ± 0,14
Omt_3	NDP CE EV CI QHTGDV FAC LFACQ	80 ± 10	37 ± 4.7
Omt_4	NDP CE EV CI QHTGDV KF CE LF CQ	40 ± 8	18.4 ± 3.3
Omt_5	NDP CE EV CI QHTGDV LAC LLACQ	~ 180 ± 1	83.7

(*) Fletterick *et. al.*; (a) not soluble in aqueous solution

As expected, none of the natural toxins (Apa-0, Scy-0, Het-0, Omt-0) showed any binding to the AR LBD, demonstrating that the intrinsic fold and helical character of the miniproteins alone is not sufficient for binding. Introduction of the FXXLF motif in the apamin (Apa-1 and Apa-2), however, resulted in a strong binding of these miniproteins to the AR with a K_i of 4.4 μM. Apa-1, containing the original glutamine in position 17, seems to be flexible enough to avoid the predicted clashes with the AR LBD surface facilitating the formation of the charge clamp (Figure 4a).

Incorporation of the same FXXLF motif in the scyllatoxin scaffold, on the other hand, only resulted in marginal binding of the resulting miniprotein Scy-2 (assumed K_i ~163 μM). Re-examination of the overlay of Scy-2 with the AR showed that, although the miniprotein does not feature any steric clashes with the AR surface, the helix of the miniprotein is too long to allow successful formation of the charge clamp (Figure 4b). The helix features around 3.6 turns and is significantly longer than the helix of Apa-1, consisting of 2.5 turns. The

miniprotein Scy-1 could not be used for affinity experiments due to its bad solubility in aqueous solutions.

Introduction of the FXXLF motif in κ -hefutoxin yielded a miniprotein (Het-1) with a moderate binding affinity to the AR. The overlay of this miniprotein with the crystal structure of the AR LBD – cofactor peptide complex indicated that the N-terminus of the miniprotein possibly collided with the AR surface via the charge clamp residue Glu₈₉₇. Interestingly, an overlay of Het-2 with the AR surface predicted the same clash, but however resulted in a high affinity for the AR. Truncated variants (Het-5 and Het-6) of these miniproteins were designed and synthesized that did not show this collision *in silico* and featured a helix length of only 1.9 turns (Figure 4c). Het-5 indeed featured a 10-fold higher binding affinity of 2.3 μ M, compared to Het-1. In contrast, the truncated miniprotein Het-6 displayed a decreased binding affinity for the AR, when compared with Het-2. Maybe the longer helix in Het-2 was flexible enough to prevent the predicted clash with the AR surface. Reduction of the helix destabilized the binding of Het-6. κ -hefutoxin features two stable helical segments. Introduction of the FXXLF motif in the second helix and a point mutation in this helix (Glu15Arg) yielded a miniprotein (Het-3) with a significant AR affinity (18.4 μ M) even though docking indicated that Asp₁₄ clashes seriously with the AR surface. As described above, the helix seems to be flexible enough to avoid the indicated clash with the AR. Shifting the FXXLF motif in Het-3 to position +1 resulted in a strong decrease in AR affinity (Het-4). The docking studies already indicated that the loop, connecting the helices of κ -hefutoxin, completely collides with the AR surface.

Next, it was focused on the mutated Om-toxin 3, which also featured two stabilized helices. The FXXLF motif was first introduced in the front part of the first helix and computational *in silico* modeling predicted that the C-terminus of the first helix is in close proximity to Lys₇₂₀. In order to avoid a contact of the miniprotein in this area a His11Gly mutation was incorporated. Electrostatic repulsion with the charge clamp residue Glu₈₉₇ was prevented via an Asp to Arg point mutation in the second amino acid of the miniprotein. The resulting miniprotein Omt-1 featured a remarkable binding affinity of 0.9 μ M. In analogy with Het-5 the FXXLF motif was also introduced in the second half of the first helix of the Om-toxin and the N-terminus was truncated by three amino acids to avoid collision with the AR surface, again featuring a helix consisting of only 1.9 turns (Figure 4). This resulted, together with a Glu to Arg mutation at position 2 to avoid electrostatic repulsion, in miniprotein Omt-2, which featured a K_I of 0.48 μ M and is the best AR binder in the library of 16 measured miniproteins. Clearly, a short stable helical segment is favorable for binding to

the AR. Introduction of the FXXLF motif in the second helix of the Om-toxin, in analogy with Het-3 and Het-4, resulted in miniprotein Omt-3 and Omt-4. Both miniproteins feature a good overlay with the AR surface, although Omt-4 resulted in a 2-fold higher affinity for the AR LBD. In Omt-4 the helix is arranged closer to the AR surface resulting in a K_I similar to Het-3 (18 μ M).

A selected set of miniproteins was further used to introduce an LXXLL motif instead of the FXXLF motif. The corresponding miniproteins Scy-3 and Omt-5 resulted in a strong decrease in binding affinity for the AR. Despite the overall similarity in peptide binding modes, the DHT-bound AR LBD is proposed to bind LXXLL motifs weakly and, instead, bind preferentially to aromatic-rich motifs that are found within the AR N-terminal AF-1 and AR-specific coactivators. Interestingly, changing the FXXLF motif in the apamin mutant Apa-1 to an LXXLL motif resulted in a high affinity for the AR LBD that was, with a K_I of 3.9 μ M, comparable to the affinity of the apamin mutants featuring an FXXLF motif. The ligand-bound AR forms a deep narrow hydrophobic groove in the LBD, which could easily accommodate the bulky side chain of phenylalanine residues that are further stabilized by the charged clamp residues^[29]. However, several cofactor peptides, bearing an LXXLL motif, were identified to bind to the AR LBD. These cofactor peptides fit loosely into the groove and the peptide backbone only forms a hydrogen bond with the Lys₇₂₀. A shift in the peptide position prevents the direct hydrogen bonding with Glu₈₉₇, which could explain the relatively lower affinity of AR for these LXXLL motifs. Possibly, the recognition motif in the apamin mutants is not alone crucial for AR binding. In Apa-3, like in Apa-1 and Apa-2, neighboring amino acids as well as the helical fold of these miniproteins might provide support for the binding stability afforded by the charge clamp residue Lys₇₂₀, thus enabling high affinity despite of using an LXXLL motif. On the other hand, the miniproteins Scy-2 and Omt-3, although featuring FXXLF motifs bind the AR only moderate (K_I : 163 μ M and 37 μ M) indicating a general destabilization by a differentiated helical fold or nearby residues. In particular Scy-2 clashes with Lys₇₂₀, most probably avoiding the formation of the charge clamp that is especially important for the interaction with peptides bearing an LXXLL motif.

In order to confirm the importance of the stabilized helical fold of the miniproteins for the AR binding affinity, mutations of the six best AR binding miniproteins were prepared. In these mutants the cysteines involved in the disulfide bridges were substituted by methionines. Thereby, the stabilization of the helical secondary structure by disulfide formation is excluded and the affinity would rely uniquely on the presence of the FXXLF motif and the surrounding amino acids. Synthesis of these methionine mutants was performed via standard

SPPS. The crude peptides were purified using preparative RP-HPLC and the purity (>95%) was confirmed by analytical RP-HPLC.

CD measurements and competitive FP experiments were performed with the synthesized methionine mutant miniproteins to investigate a change in their secondary fold and a possible resulting change in the AR binding affinity (Figure 11). For Apa-2 Met and Apa-3 Met there is still a low degree of α -helicity observable, with the typical maximum at 192 nm and both minima at 208 nm and 222 nm, suggesting partial folding of these miniproteins (Figure 11a). The binding affinity of these miniproteins for the AR is however up to 100 times decreased (Figure 11b; Table 3). This result indicates the necessity of the stabilized helical fold of the miniprotein to highly improve the binding affinity for the AR. Het-5 Met and Omt-1 Met feature the characteristic minima of random coil structures^[52] with a negative band at 203 nm and a large shoulder in the 215-220 nm range of relatively strong dichroic intensities. Similar CD spectra could be observed for miniproteins in which the disulfide bridges are reduced^[56, 69]. Interestingly, despite of the largely unfolded structure, both miniproteins featured only a small decrease in AR binding compared to their cysteine analogues (Figure 11b; Table 3). Possibly, these miniproteins fold into a partial helical segment upon binding to the AR LBD surface, similar to normal linear peptides, and the amino acid sequence surrounding the FXXLF motif is highly favoured for this and for the interaction with the AR surface. The introduction of methionines in the miniproteins Het-2 Met and Omt-2 Met as well led to significant changes in the CD spectrum, demonstrating the complete loss of the α -helical character of these miniproteins (Figure 11a). This was accompanied by an up to 100 fold decrease in AR binding affinity for Het-2 and Omt-2 (Figure 11b; Table 3).

All methionine mutated miniproteins featured a significant decrease in α -helicity suggesting that the helical structures of the oxidized miniproteins with cysteines were dependent on the disulfide bridges (Table 3). Exchanging the cysteines, responsible for the formation of disulfide bridges, by methionines led to a significant change in the miniprotein structure. Without the correct α -helical fold, these miniproteins show only a reduced binding affinity for the hydrophobic groove on the surface of the AR. Further, it has to be mention that the sequences surrounding the FXXLF motif apparently indicate a high importance of the pre-folded structures in terms of affinity towards the AR LBD cofactor groove, while the surrounding sequences in well-folded miniproteins is only of secondary importance.

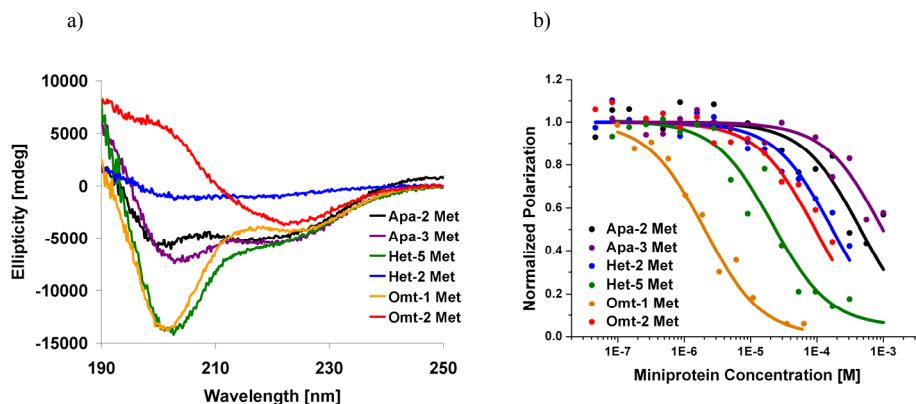


Figure 11: (a) Far UV circular dichroism spectra of 50 μM methionine mutated miniproteins recorded in 50 mM potassium buffer pH 7.4. The CD spectra revealed a decrease of α -helicity for these miniproteins. (b) Normalized competitive FP displacement data of the AR LBD – Fl-FXXLF peptide by miniproteins containing methionines instead of cysteines resulting in a decreased affinity.

Table 3: Miniproteins bearing methionines instead of cysteines and their binding affinity for the AR LBD

Name	Sequence	IC ₅₀ (μM)	K _I (μM)
Apa-2 Met	MNMKAPETA <u>F</u> M <u>LF</u> MQGH	$\sim 450 \pm 160$	$\sim 209 \pm 72$
Apa-3 Met	MNMKAPETALM <u>LL</u> MQGH	$\sim 980 \pm 210$	$\sim 457 \pm 98$
Het-2 Met	GHAM <u>YR</u> <u>F</u> M <u>LF</u> GNDEETM <u>KERM</u>	$\sim 170 \pm 50$	$\sim 79 \pm 23$
Het-5 Met	AM <u>YF</u> <u>N</u> <u>LF</u> EGNDEETM <u>KERM</u>	20 ± 10	9 ± 4.3
Omt-1 Met	NR <u>F</u> <u>M</u> <u>LF</u> MIQGTGDVKAM <u>EEAM</u> Q	2 ± 1	0.63 ± 0.16
Omt-2 Met	M <u>R</u> <u>F</u> <u>V</u> <u>ML</u> FHTGDVKAM <u>EEAM</u> Q	90 ± 40	42 ± 18

^a Introduced methionines are highlighted in grey, inserted mutations are underlined and the signature amino acids are represented in **bold**.

3.4 Co-Crystallization Approach of the Miniproteins in Complex with the AR LBD

In order to evaluate the exact molecular interactions of the miniproteins with the AR LBD, crystallization studies were performed. X-ray co-crystallization allows the analysis of the three-dimensional composition of each atom in a protein crystal via a generated diffraction pattern^[70]. For the generation of suitable protein crystals, the AR LBD was expressed in the presence of an excess of the agonist DHT and purified with an N-terminal His₆-tag using immobilized metal ion affinity chromatography (IMAC). After cleavage of the His₆-tag and intensive dialysis, the purity of the collected AR LBD fractions was analyzed via SDS-PAGE (Figure 12). Due to protein aggregation problems, the AR LBD could only be concentrated to 2-4 mg/ml, resulting in 66-132 μM protein solutions.

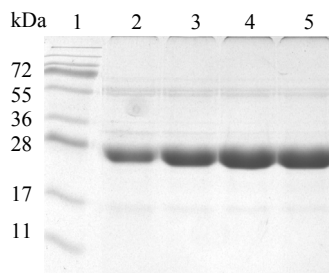


Figure 12: Analysis of the purification of the AR LBD (aa 664 - 919) cleaved from the N-terminal His₆-tag via SDS-PAGE (15% gel; Coomassie stained). Lane 1: Molecular weight marker, lanes 2-5: AR LBD (calcd. mass: 30245.4 Da).

For crystallization, the miniproteins must be highly soluble in the protein buffer and should feature a high affinity for the AR. Therefore, the miniproteins Apa-3 and Het-5, achieving these characteristics, were chosen for initial crystallization experiments. After dialysis against protein buffer, each miniprotein was incubated with the DHT occupied AR LBD in miniprotein to AR LBD ratios of 2:1, 3:1, 4:1 and 5:1, respectively. As a control, a protein approach containing the reference FXXLF peptide was crystallized in parallel. Subsequent crystallization of the AR LBD/DHT/peptide complex was carried out using the sitting drop vapor diffusion method with commercial screening solutions. After one to two days, small, but promising rod shaped and cylindrical crystals appeared for several conditions (Figure 13). These crystals grew to a final size of 90 μm x 13 μm x 10 μm within 35 days.

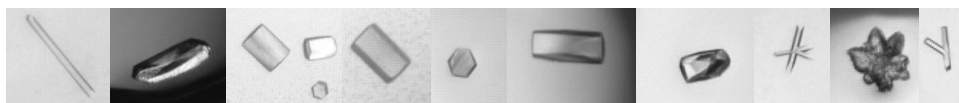


Figure 13: Observed crystals from the AR LBD/DHT/miniprotein screening at 4 °C and 18 °C in sitting drops using commercial screening solutions. The general size of each crystal was 90 μm x 13 μm x 10 μm .

Using the sitting drop method at 4 °C, these crystals could be reproduced for large scale and fine screening experiments. However, despite many variations in the composition of mother liquor, drop size, protein-miniprotein ratio and incubation at 18 °C and 4 °C, no increase in crystal size could be observed within the next 35 days. As the AR LBD after purification via IMAC and cleavage of the His₆-tag resulted in large aggregations, it was decided to express and purify the AR LBD as a fusion protein with GST. After purification via glutathione affinity chromatography and cleavage of the GST moiety, the AR LBD was further purified via cation exchange chromatography (Figure 7). The protein could be

concentrated subsequently to 6 mg/ml before aggregating. Despite the higher concentration no improvement in crystal size could be detected for the complex of AR LBD, DHT and miniprotein or reference peptide.

Although promising crystals for the AR/DHT/miniprotein screening could be detected, the low protein stability and accordingly the relative low protein concentration are most probably the main reasons for the small size of these crystals. Trials to improve the stability of the AR LBD would facilitate a higher concentration of protein, and thus the formation of high quality crystals. Further, it has to be kept in mind that the AR LBD needs reducing conditions for stability, while the miniproteins need to be oxidized to form the stabilizing disulfide bridges. Further protein conditions have to be screened to identify a suitable condition for both the AR LBD and the miniprotein.

3.5 Estrogen Receptor Binding Studies of the Miniproteins

Despite the overall similarity in the coactivator binding modes, the DHT-bound AR LBD prefers the binding of the FXXLF motif, while the estradiol (E₂)-bound ER LBD prefers the LXXLL motif. Both nuclear receptors form a hydrophobic groove on the surface of their LBD at a region that is located between two residues that form a charge clamp to stabilize the position of the cofactor helix. However, the hydrophobic groove of the AR is deeper and prefers to accommodate the bulky side chain of phenylalanine residues^[29]. The cofactor binding groove of the E₂-bound ER is more shallow and suited for the accommodation of the leucine residues^[3]. In order to investigate the selectivity of the designed miniproteins for the AR, the binding affinity of these miniproteins for the estradiol-bound ER LBD was investigated. Therefore, both the ER α and the ER β LBD were expressed and purified with an N-terminal His₆-tag using IMAC. The purities of the His₆-ER α LBD and the His₆-ER β LBD were confirmed by SDS-PAGE (Figure 14).

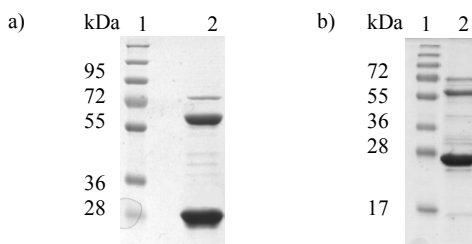


Figure 14: Analysis of the purification of (a) the His₆-ER α LBD₃₀₂₋₅₅₃ (calcd. mass: 31059.4 Da) and (b) the His₆-ER β LBD₂₆₀₋₅₀₂ (calcd. mass: 27534.7 Da) via SDS-PAGE (15%; Coomassie stained). Lane 1: molecular weight marker, lane 2: purified ER LBD. The impurity with a size of 60 kDa is most probably the bacterial heat shock protein GroEL that strongly binds the nuclear receptor and could not be removed during the purification.

Subsequently, a fluorescence polarization (FP) assay, similar to the one developed for the AR, was set up for both ER isoforms. Therefore, a peptide was synthesized (NH₂-LTERHKILHRLLLQEGSPSD-CO₂H), bearing the sequence of the second LXXLL motif in the steroid receptor coactivator 1 (SRC 1 Box 2). This peptide is known to bind both ERs with high affinity. The same peptide was synthesized with an N-terminal cysteine in order to specifically label the peptide with fluorescein. The affinity of the LXXLL peptide for the E₂-occupied ER α and ER β LBD was measured using FP by varying the concentration of the ERs in the presence of 0.1 μ M fluorescein-labelled peptide and saturating concentrations of E₂ (Figure 15a,b). The K_D values, 0.77 μ M for the ER α and 0.3 μ M for the ER β LBD, were determined from three replicates of each experiment. These values were in agreement with previously measured K_D values of this peptide binding to the ER LBD subtypes^[71, 72].

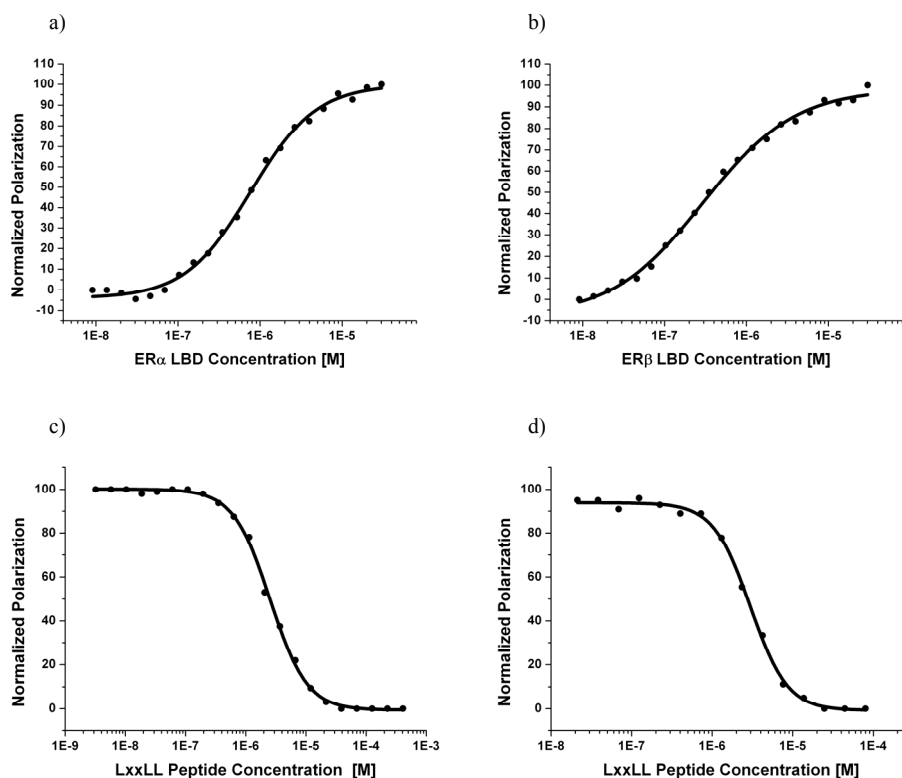


Figure 15: Titration of a constant concentration (0.1 μ M) of the fluorescein-labelled LXXLL peptide with increasing concentrations of (a) the E₂ saturated ER α LBD to estimate the K_D (0.77 μ M \pm 0.037 μ M) of the protein-peptide pair. (b) Binding affinity of the E₂-bound ER β LBD with a K_D of 0.3 μ M \pm 0.019 μ M. (c) Displacement of the fluorescein-labelled LXXLL peptide from the ER α LBD by increasing concentrations of unlabelled LXXLL peptide (IC₅₀: 2.54 μ M, K_I: 0.8 μ M). (d) Competitive titration analysis of the ER β LBD – Fl-LXXLL peptide complex (1 μ M : 0.1 μ M) by free LXXLL peptide (IC₅₀: 2.98 μ M; K_I: 0.51 μ M).

In the subsequent competitive displacement FP studies, the concentrations of receptor and fluorescent peptide are held constant. The condition chosen for the ER α LBD was 1 μ M. At this concentration 70% of the labelled LXXLL peptide exists as a complex with the ER α LBD. As the ER β LBD featured a lower K_D of 0.3 μ M for the LXXLL peptide binding, a concentration of 0.5 μ M was chosen, at which around 80% of the labelled peptide exists as AR LBD – FI-LXXLL peptide complex. The competitive displacement FP led to an calculated K_I of 0.8 μ M for the ER α LBD and a K_I of 0.51 μ M for the ER β LBD (Figure 15c,d). The successful established competitive displacement FP assay was used to screen a selected set of the folded miniproteins for potential ER LBD binding (Figure 16).

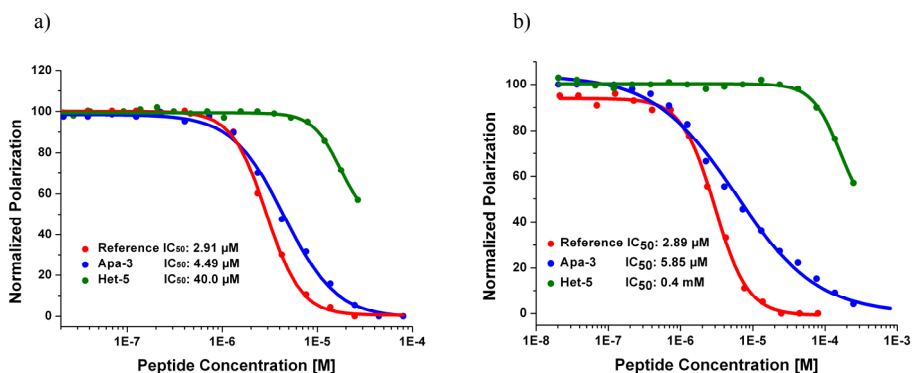


Figure 16: Competitive displacement FP analysis of (a) the ER α LBD – FI-LXXLL peptide complex and (b) the ER β LBD – FI-LXXLL peptide complex by increasing concentrations of two miniproteins (Apa-3 in blue and Het-5 in green) and the reference LXXLL peptide (red).

Most of the miniproteins displayed no binding to the ER LBDs as expected. All miniproteins were *in silico* designed to interact with the narrow deep hydrophobic groove on the surface of the AR LBD. Especially the introduction of the bulky phenylalanine residues as part of the AR specific FXXLF motif prevents the binding of the miniprotein in the shallow coactivator recognition groove in the ER LBD. Further, the position of the FXXLF motif within the miniprotein helix and the length of the helix were optimized for fitting to the AR surface between the charge clamp residues. These results demonstrate the selectivity of the designed miniproteins for the AR.

Table 4: Binding affinities of the miniproteins for the ER α and ER β LBD estimated by FP

Name	ER α LBD	ER α LBD	ER β LBD	ER β LBD
	IC ₅₀ [μ M]	K ₁ [μ M]	IC ₅₀ [μ M]	K ₁ [μ M]
Reference	2.91	0.96	2.89	0.49
LXXLL				
Apa-1	n.a.	n.a.	n.a.	n.a.
Apa-2	n.a.	n.a.	n.a.	n.a.
Apa-3	4.49	1.64	5.85	1.17
Het-2	n.a.	n.a.	n.a.	n.a.
Het-5	40	16.92	410	94.69
Omt-2	n.a.	n.a.	n.a.	n.a.
Omt-5	n.a.	n.a.	n.a.	n.a.

n.a. not binding affinity could be detected

Interestingly, two miniproteins nevertheless featured a remarkable binding for both the ER α and the ER β LBD (Table 4). One of these miniproteins is Apa-3. This miniprotein features an LXXLL motif in its helix and is therefore a suitable candidate to bind to the shallow groove of the ER LBD. The binding affinity of Apa-3 for the AR was determined at 3.9 μ M. For both ER LBDs a similar K₁ could be calculated (Figure 16; Table 4). Interestingly, Omt-5, the second measured miniprotein with an LXXLL motif, showed no interaction with the ER LBDs. Comparisons with the AR showed only a moderate binding affinity to the AR. Most likely, other reasons than the introduced binding motif are responsible for the bad binding affinity of Omt-5 for the AR and ER LBDs. Het-5, featuring a K₁ of 2.3 μ M for the AR, binds with a moderate affinity of 17 μ M to the ER α LBD and with a K₁ of 95 μ M to the ER β LBD. Although the K₁ is 7-fold and 40-fold lower than for the AR, this result is astonishing as Het-5 features an FXXLF motif in its helix instead of an LXXLL motif.

A further ER binding experiment was performed in order to confirm the observed results. This fluorescence resonance energy transfer (FRET) assay measured the binding affinity of the miniproteins for the ER LBDs in dependency of the agonist E₂. In the used experiment, developed by Organon, now Merck, a complex of allophycocyanine (APC) conjugated streptavidin and biotinylated LXXLL peptide must be recruited by a complex of His₆-tagged ER LBD and europium (Eu) labelled α -His antibody, to bring both fluorophores close to each other (Figure 17).

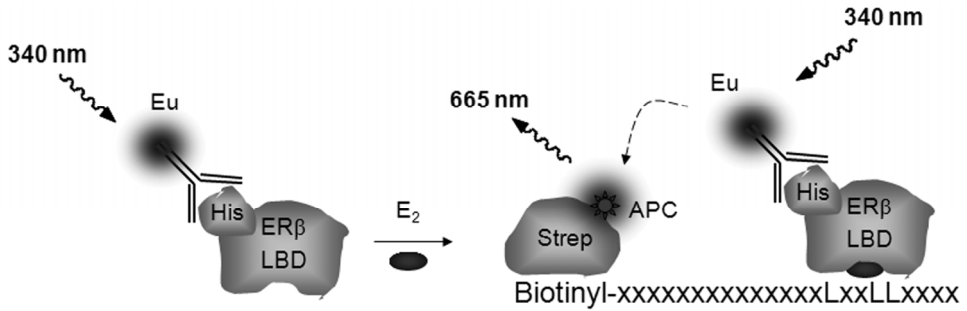


Figure 17: Illustration of the *in vitro* cofactor recruitment FRET assay. Binding of an agonist to the His₆-tagged ERβ LBD will recruit a biotinylated coactivator peptide. Eu-labeled α-His antibody and streptavidin-conjugated APC will assemble into the complex resulting in FRET. The amino acid sequence of the peptide used in this experiment belongs to the fourth LXXLL motif in the SRC 1a.

The recruitment of the LXXLL peptide by the ER LBDs is induced by increasing concentrations of E₂. Competitive displacement of the LXXLL peptide by different concentrations of miniprotein resulted in a horizontal decrease of the fluorescence signal without any loss in the binding affinity of E₂ (Figure 18). Although this experiment allows no determination of a K_I for miniprotein binding to the ER, the significant decrease in the fluorescent signal for Apa-3 and Het-5 confirmed the results observed in the FP assay. While the lowest concentration of 0.1 μM Apa-3 already caused a signal loss for both the ERα and the ERβ LBD, the FRET signal for the LXXLL peptide-bound ER could only slightly be decreased by high concentrations (10 μM) of the miniprotein Het-5. The other measured miniproteins did not display any effect on the FRET signal. These results again demonstrate the high selectivity of the designed miniproteins for the AR. This selectivity mainly depended on the introduction of the FXXLF motif in the helix of the miniprotein.

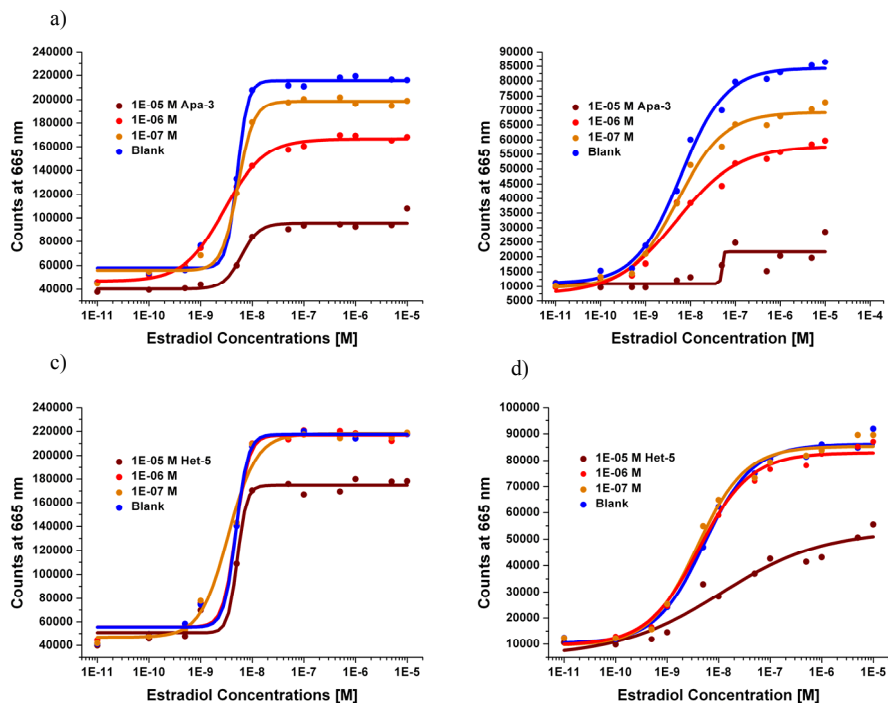


Figure 18: Influence of the mini-proteins Apa-3 and Het-5 on the E_2 -induced peptide recruitment by the ER α (a and c) and the ER β LBD (b and d) using FRET. Concentrations of the ER LBD: 100 nM and of the LXXLL peptide: 10 nM. Increasing concentrations of mini-protein lead to a decrease in the FRET signal due to competitive LXXLL peptide displacement.

In order to get more information about the ER binding process of the mini-protein Apa-3, an isothermal titration calorimetry (ITC) assay was set up. Beside the quantification of binding affinities, ITC allows the determination of further thermodynamic parameters of the interaction. The binding reaction of a peptide to the ER LBD is an exothermic process that increases the temperature in the analyzed system. ITC is a thermodynamic technique that directly measures the heat released during this molecular binding event. Measurement of this heat allows accurate determination of binding constants (K_D), reaction stoichiometry (N), enthalpy (ΔH) and entropy (ΔS), thereby providing a complete thermodynamic profile of the molecular interaction in a single experiment. The concentration of the protein is suggested to be at least 10-fold higher than the K_D of the binding reaction, while the concentration of the titrating peptide should be 10-fold higher than the protein concentration. Here, an ER LBD concentration of 20 μ M and an Apa-3 concentration of 0.2 mM were used. Each experiment was performed in the presence of E_2 and in triplicate. For comparison, the reference LXXLL peptide is used in a separate ITC experiment. The integrated heat data have been corrected for

heats of peptide dilutions and buffer effects, which are relatively small for the chosen conditions (Figure 19). Each exothermic heat pulse (upper panel) corresponds to the injection of 8 μ l of peptide into the solution of ER LBD. The integrated heat data (lower panel) constitute a differential binding curve that fit to a standard single-site binding model to give, in this instance, the stoichiometry (N), the affinity (K_A) and the change in enthalpy (ΔH) of peptide binding (Table 5). The K_D of peptide binding was calculated from $1/K_A$.

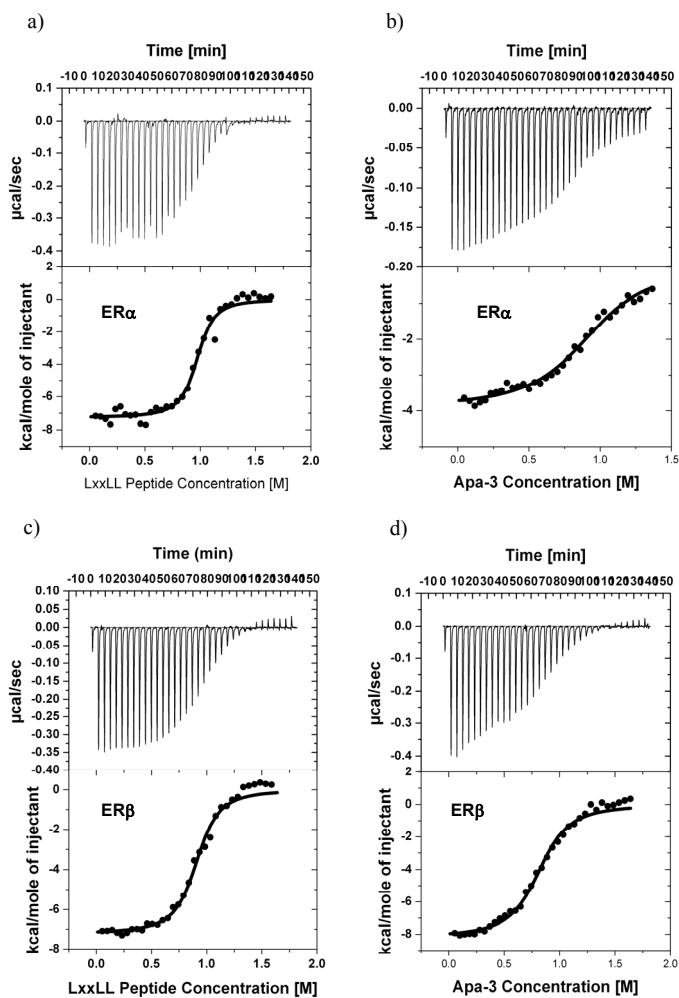


Figure 19: Representative ITC titrations of the LXXLL peptide and Apa-3 into the ER α (a,b) and the ER β LBD (c,d) in the presence of E₂. Titrations were performed in PBS buffer at pH 7.5 at 26 °C. Presented are the heat pulses obtained for 35 automatic injections, each of 8 μ l, of 0.2 mM control peptide or Apa-3 solution into the sample cell containing the ER LBD solution at a concentration of 20 μ M. The integrated heat data were fitted to a single-site binding model to determine the thermodynamic parameters of binding.

Table 5: Binding stoichiometries and thermodynamic parameters of ER α/β - peptide interactions in PBS buffer (pH 7.4) at 26 °C using ITC

Ligand	K_A $10^6 M^{-1}$	K_D μM	ΔG kcal/mol	ΔH kcal/mol	$T\Delta S$ kcal/mol	N
ER α LBD						
Reference	6.70 \pm 1.59	0.15 \pm 0.03	-7.45 \pm 0.12	-7.27 \pm 0.12	0.18	0.96 \pm 0.01
Apa-3	0.66 \pm 0.08	1.46 \pm 0.11	-4.26 \pm 0.06	-3.91 \pm 0.06	0.353	0.96 \pm 0.01
ER β LBD						
Reference	1.74 \pm 0.22	0.57 \pm 0.06	-8.21 \pm 0.12	-8.18 \pm 0.12	0.031	0.82 \pm 0.009
Apa-3	1.21 \pm 0.15	0.83 \pm 0.09	-5.92 \pm 0.15	-5.69 \pm 0.08	0.226	1.24 \pm 0.01

The stoichiometry (N) of both the ER α and the ER β LBD binding by the reference LXXLL peptide and Apa-3, respectively was close to 1, which means that the apparent stoichiometry for the interaction is one peptide molecule per ER LBD molecule. Apa-3, as well as the reference LXXLL peptide, binds only to the hydrophobic groove of the ER and not to a second binding site within the LBD as expected. The affinity (K_D) of Apa-3 for ER α was only 10-fold lower (1.46 μM) compared to the natural LXXLL peptide (0.15 μM). Similar results were found for the ER β recruitment of the peptides. Again the natural peptide featured only a slightly stronger affinity for the receptor (0.57 μM) compared to Apa-3 (0.83 μM). Comparisons of the K_D values with the results obtained from the FP showed that the affinities were in the same region and confirmed the successful establishment of ITC to determine binding affinities of miniproteins for the ER.

The binding interactions between the peptide or the miniprotein and the ER LBDs are driven by both favorable enthalpic (negative ΔH) and entropic (positive $T\Delta S$) interactions. The favorable enthalpic interactions most likely reflect the contributions from charged residues on the ER surface. In both cases, the peptide binds to the hydrophobic groove on the surface of the ER LBD. The peptide is held in place through the interactions of its two leucine residues with the hydrophobic groove constituents, but also by hydrogen bonds that involve two conserved amino acids of NR LBDs. Both residues are hydrogen-bonded to a main-chain peptide bond of the LXXLL motif and together form a “charge clamp” that stabilizes the peptide - receptor interaction. ITC measures the global thermodynamic parameters of a system so the ΔH value is the result of the sum of these hydrogen bonds, changes in protein – solvent and peptide – solvent interactions, and changes in bonding within the protein and peptide also contributes. Beside these H bonding interactions, the interface of the peptide and ER is predominantly apolar. The burial of such exposed hydrophobic faces results in favorable entropy (positive $T\Delta S$).

The data confirm that the reference LXXLL peptide and the miniprotein Apa-3 bind to only one position within the ER LBD and that the amino acids recognized in this binding pocket are the same for both peptides. But it should be noted, that the enthalpy and entropy changes reported in Table 5 are apparent values observed in a specific buffer at a single temperature. From these data the role of protonation in the recognition of the reference peptide or Apa-3 binding to the ER α and the ER β LBD can not be evaluated. Ideally, multiple temperatures and buffer systems should be examined but difficulty in obtaining the quantity of peptide material prevented exhaustive analyses.

3.6 Co-Crystallization Approach of the Miniprotein Apa-3 in Complex with the ER β LBD

In order to get deeper insights in the exact binding mechanism of the miniproteins on the molecular level, x-ray crystallography was envisioned to solve the crystal structure of a miniprotein with a NR LBD. Due to the problems in the co-crystallization of a miniprotein complex with the AR LBD (3.4), it was decided to co-crystallize the miniprotein Apa-3 in complex with the ER β LBD. The ER β LBD could be successfully crystallized in other experiments (Chapters 4 and 6) and therefore, is a promising tool for co-crystallization studies with Apa-3. The ER β LBD was expressed in the absence of E₂ in *E. coli* and purified without affinity-tags via estradiol affinity chromatography. After elution with increasing concentrations of E₂ from the affinity column, the ER β LBD was further purified using size exclusion chromatography (SEC). Subsequently, the purity of the ER β LBD was confirmed via SDS-PAGE (Figure 20).

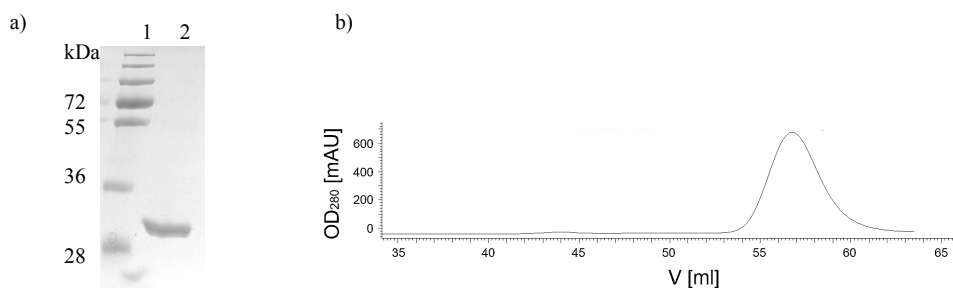


Figure 20: Analysis of the purification of the ER β LBD. (a) SDS-PAGE gel (15%; Coomassie stained). Lane 1: Molecular weight marker, lane 2: ER β LBD_{MD[D261-L500]DD} (calcd. mass: 27553.5 Da). (b) Elution profile of ER β LBD_{MD[D261-L500]DD} after SEC using FPLC.

Co-crystallization studies with the ER β LBD were performed similar to the trials used for the AR. The E₂ bound protein was concentrated to 12 mg/ml and incubated with the dialyzed Apa-3 in a molar protein to peptide ratio of 1:2 and 1:5. Commercial screening solutions were mixed 1:1 with the ER β LBD/E₂/Apa-3 complex in sitting drops and incubated at 18 °C and 4 °C for 35 days against the reservoir solution. After three days, promising small rod shaped crystals appeared in one condition at 4 °C with a final size of 100 μ m x 10 μ m x 10 μ m (Figure 21). Although the observed crystals look promising they are still too small for diffraction measurements. Intensive large scale reproduction studies have to be undertaken to increase the overall size of the crystals. Further, modification of several parameters, like pH, drop size, protein concentration and protein – miniprotein ratio could improve the quality of the detected crystals.



Figure 21: Observed crystals from the ER β LBD/E₂/Apa-3 screening at 4 °C in sitting drops using a solution containing 0.1 bicine pH 9.0, and 20% PEG 6000. The general size of each crystal was 100 μ m x 10 μ m x 10 μ m.

3.7 Conclusion

In total 17 novel miniproteins were designed and synthesized based on four different miniprotein toxins. Of 16 measured miniproteins, seven featured binding affinities between 0.5 and 5 μ M, and six between 10 and 40 μ M. Only three miniproteins (Scy-2, Scy-3 and Het-4) featured limited binding. These high affinities and hit-rates are remarkable. One of the best known peptide AR binders (NH₂-SSRFESLFFAGEKESR-CO₂H), obtained *via* phage display, had a K_1 of 3.9 μ M in the described FP assay. This phage display derived peptide resulted from a large library and has accordingly also been optimized for the amino acids surrounding the FXXLF motif, important for secondary interactions with the AR.

The 16 miniproteins described here have not been significantly optimized with respect to the amino acids surrounding the FXXLF motif, but nevertheless already feature higher binding affinities. *In silico* design thus offers a rapid entry into potent binders for the AR. This holds great promises to generate AR peptidic cofactor mimics that bind with similar affinity as those developed for the ER.

The obtained results furthermore show that apparently the presence of the FXXLF motif on a preformed stable helical segment in the miniprotein is the critical feature that conveys high binding affinity for the AR. Mutants of some miniproteins, in which the four cysteines that account for the stable fold were mutated to four methionines, typically featured a 50 to 100-fold lower AR binding affinity than the miniproteins with cysteines. These mutants further indicate that the sequence of the surrounding amino acids is apparently much more important in unstructured peptides. Furthermore, the helical segment of the peptide should be short, approximately 2 turns, as a longer helix clashes with the AR surface or prevents the formation of the charge clamp.

Introduction of the ER specific LXXLL motif instead of an FXXLF motif generally resulted in a decrease of activity. However, for one miniprotein, Apa-3, no reduction in the affinity for the AR LBD could be detected. This miniprotein, although designed to bind to the AR LBD, also displayed a high affinity for the ER α and ER β LBDs, while all other miniproteins did not featured significant binding to the ER LBD. This result confirmed the selectivity of the designed miniproteins for the AR.

The miniproteins described here now provide an entry to study the important effects of the amino acids surrounding the FXXLF motif on NR binding, disconnecting the effect of these amino acids on helix stability, as occurring in linear peptides, from surface recognition. As such, other miniprotein libraries should give even more detailed molecular insights into the recognition of NRs by coactivators and provide the molecular requirements to generate new coactivator binding inhibitors for NRs.

The initial studies to crystallize the miniproteins in complex with NR LBDs provide an entry to evaluate the exact molecular interactions of the miniproteins with the NR. Further attempts to generate miniproteins, stable under cellular conditions offer the possibility for applications in mammalian cell-based inhibition studies in the future.

3.8 Experimental Section

General Information. Rink Amide MBHA resin with loading of 0.72 mmol/g and Rink Amide AM resin LL (100-200 mesh) with loading of 0.34 mmol/g were purchased from Novabiochem. Fmoc-protected amino acids were purchased from MultiSyntech and Novabiochem in their appropriately protected form. All other reagents were purchased from Aldrich-Sigma, Fluka and Acros. LC-ESI-MS was carried out by using an Agilent 1100 series binary pump together with a reversed-phase HPLC column (Macherey-Nagel) and a Finnigan Thermoquest LCQ. If not otherwise stated, the following gradient program was used for analytical LC-MS: flow: 1 mL/min, solvent A: 0.1% HCO₂H in H₂O, solvent B: 0.1% HCO₂H in CH₃CN, A/B: 90/10 (0-1 min) to 0/100 (over 10 min), 0/100 (12 min). Purification of products by RP-HPLC was performed in an Agilent 1100 Series Purification Platform using a NUCLEODUR[®] C18 Gravity preparative column from Macherey-Nagel (21

x 250 mm) and flow rate of 25 mL/min. The products were eluted by using different solvent gradients of solvents A and B (solvent A = 0.1% TFA/H₂O; solvent B = 0.1% TFA/CH₃CN). UV signal at 210 nm was used for detection.

Folded miniprotein synthesis.

The *in silico* design of all miniproteins was performed by Dr. S. Folkertsma (Organon,) and Dr. B. Vaz.

A. Linear miniprotein synthesis (automated). The synthesis, folding and purification of all miniproteins was done by Dr. B. Vaz. All sequences were synthesized from C- to N-terminus on solid support^[45], using an automatic solid phase synthesizer (Syro II, Multisynthec) on a 72 μmol scale (100 mg of resin, loading of 0.72 mmol/g). The coupling of amino acids was carried out following the standard Fmoc-chemistry^[46], using HOBt/DIC^[47, 73] (4 equiv.) as amino acid activation, DMF as solvent and 4 equiv. of the protected Fmoc-amino acids. The resin was first swollen with DMF (1 x 20 min) and the Fmoc protecting group was removed by treatment with piperidine/DMF (2/3, 1 x 3 min; 1/4 1 x 10 min), then washed with DMF (6 x 1 min). One cycle of peptide elongation consisted of the following steps. First, the deprotected resin was treated for 50 min with a cocktail containing the appropriate amino acid (4 equiv, solution 0.3 M in DMF) with equimolar addition of HOBt and DIC (4 equiv, solution 0.3 M in DMF). After washing the resin with DMF (4 x 1 min), the Fmoc protecting group was removed by treatment with piperidine/DMF (2/3, 1 x 3 min; 1/4, 1 x 10 min). After deprotection, the resin was again washed with DMF (6 x 1 min). These steps were repeated until the miniprotein sequence was complete. After the completion of the sequence, the resin was subsequently washed with DMF (5 x 30 s), CH₂Cl₂ (5 x 30 s) and Et₂O (5 x 30 s) and dried under vacuum for 2-3 h. Cleavage and side chain deprotection was carried out by treatment of the resin for 3 h with a cleavage cocktail containing TFA/H₂O/EDT/TIS (94:2.5:2.5:1)^[48, 49]. The cleaved resin was washed with TFA (3 x 15 s) and the cleaved linear miniprotein was collected, concentrated by rotary co-evaporation with toluene into less than 1 mL solution and precipitated by addition of cold Et₂O (10 mL). The mixture was cooled in a liquid N₂ bath for 1 min, centrifuged (4000 rpm, 5 min, 4 °C) and the Et₂O was decanted from the pellet. Cold Et₂O was added again and the procedure was repeated twice. The crude peptide obtained was dissolved in H₂O/CH₃CN and lyophilized to dryness.

B. Folding and purification of miniproteins. The crude products obtained from lyophilization were dissolved in a mixture 2:1 of sodium phosphate buffer (100 mM NaH₂PO₄/Na₂HPO₄, 1:17.8, pH 8) and trifluoroethanol (TFE) to a final concentration of 0.5 mg/mL. In order to reduce previously formed random disulfide bridges, the solution was stirred with 2 equiv. of TCEP for 2 h at room temperature. After addition of 10 vol% of dimethylsulfoxide (DMSO), subsequent oxidation was carried out by exposure to air for 24-48 h. The transformation was followed by LC-MS. When the oxidation was complete, the TFE was removed by evaporation and the resultant aqueous solution was lyophilized. The crude product was dissolved in water and the salts were removed by filtration through a SepVak® Vac C18 column. After elution of the salts with 2 volumes of water, the oxidized miniprotein was eluted with a mixture of H₂O/CH₃CH (1:1) and then pure CH₃CN. Fractions containing acetonitrile were combined and lyophilized. The crude oxidized miniproteins were purified by reverse-phase HPLC on a Nucleodur C18 Gravity column (125 x 21 mm, Macherey-Nagel) with a linear gradient of A (0.1% TFA in H₂O) and B (0.1% TFA in CH₃CN) from 10% of A to 40%-60% of B and flow rate of 25 mL min⁻¹ and were detected at 210 nm using a diode array UV/VIS detector. The identities and purities of the purified folded miniproteins were assessed by LC-MS (ESI mass spectrometry). The correct

folding of the miniprotein was assessed by comparison of their CD spectra with those of the natural miniprotein. Following purification, all miniproteins were lyophilized and kept at $-20\text{ }^{\circ}\text{C}$.

Synthesis and fluorescent labeling of the reference peptides. A high affinity peptide for the androgen receptor was obtained by Fletterick *et al.*^[30] by phage display (SSRFESLFAGEKESR). This peptide was selected for the competitive displacement fluorescence polarization assay as the competitive cofactor binding inhibitor (CBI) for the AR LBD. For the competitive displacement fluorescence polarization assay using the ER LBD a peptide from SRC-1 Box 2 (⁶⁸⁵LTERHKILHRLQLQEGSPSD⁷⁰³) was synthesized. The synthesis of both peptides was carried out on solid phase using an automated solid phase synthesizer (see above). Analogues of these peptides featuring an N-terminal cysteine, reactive towards a fluorescein label (FI-CSSRFESLFAGEKESR and FI-CLTERHKILHRLQLQEGSPSD), were synthesized as well. To a solution of the peptide in potassium phosphate buffer (100 mM, pH 7.2, previously degassed by sonication for 1h at rt, 1mg/mL) TCEP-HCl (10 equiv.) was added. The resulting clear solution was stirred for 1h at rt. Subsequently, a solution of 4(5)-(Iodoacetamido)fluorescein (5 equiv.) in DMSO (10 mg/mL) was added and the homogenous mixture was stirred at rt. The course of the reaction was followed by LC-MS. After completion of the reaction, the solution was quenched with ethanethiol. The resulting mixture was directly lyophilized. The labeled peptides were purified by RP-HPLC (NUCLEODUR® C18 Gravity column, Macherey-Nagel). The eluents used were A: H₂O (+0.1% HCOOH) and B: CH₃CN (+0.1% HCOOH). The method featured a 20 min gradient of a mixture A/B 90:10 to A/B 50:50 affording the expected fluorescein labeled peptide with high purity (>99 %) and 10% yield for the labeling process. The concentration of the solution of the FI-peptides in HEPES buffer was determined using the UV absorption at 495 nm (molar absorption of fluorescein 75000 M⁻¹ cm⁻¹) with four different dilutions of the original FI-peptide solution, measuring the absorption value at 495 nm (the zero values were corrected by the corresponding values measured at 580 nm).

Circular Dichroism. The CD measurements of all miniproteins were done by Dr. B. Vaz. The CD spectra were measured at 20 °C using a 50 μM peptide solution in a 5 mM potassium phosphate buffer (pH 7.4). An absorption spectrum from 185 nm to 280 nm was recorded using a UV-Spectrometer (Jasco, J-815) with a JASCO PTC-425S temperature controller. Shown measurements are averages of ten scans using 5 mM potassium phosphate buffer as reference. Using 1 mm path-length Quartz cells (Hellma, 110-QS) a stopped scan was employed with a 0.5 s response and a 2 nm bandwidth, and a data pitch of 0.2 nm. Mean residue ellipticity $[\theta]_R$ in [deg⁻³ L mol⁻¹ cm⁻¹] was calculated using equation 1^[56]:

$$[\theta]_R = \frac{[\theta]}{10 \times (n-1) \times c \times l} \left[\frac{\text{deg}^{-3} \times L}{\text{mol} \times \text{cm}} \right] \quad (\text{Eq. 1})$$

where n is the number of amide bonds, c is the concentration (mol/L), l is the pathway length (cm), $[\theta]$ is the measured ellipticity (deg⁻³) and $[\theta]_R$ is the mean residue ellipticity.

Polymerase Chain Reaction. PCR^[74] was performed using a T3000 thermocycler (Biometra) or a Mastercycler epgradient (Eppendorf). The general PCR preparation was consisted of a total volume of 50 μl containing 1 x *Pfu* DNA polymerase reaction buffer (Fermentas), 5-50 ng dsDNA template, 125 ng of each primer (MWG

Biotech) and 200 μ M dNTP mix (Fermentas) and 2.5 U/ μ l *Pfu* DNA polymerase (Fermentas). The amplification of DNA resulted from an initial denaturation at 95 °C for 5 min followed by 30-40 cycles of 30 s denaturation at 95 °C, 45 s of hybridization at specific temperatures (5°C lower than melting temperatures of the primers) and DNA synthesis for 2 min at 72 °C. The last step was a terminal DNA synthesis for 5 min at 72 °C. Control and purification of the amplification product was performed by agarose gel electrophoresis (0.8% agarose) in 40 mM Tris/Acetate pH 7.6, 1 mM EDTA, and SYBR Safe DNA gel stain (Invitrogen). Isolation of DNA from the gel was performed using the Qiaquick PCR purification kit or the Qiaquick gel extraction kit (Qiagen). Isolation of plasmid DNA from *E. coli* was performed using the QIAprep Spin Miniprep Kit (Qiagen).

Plasmids. The hAR LBD (residues 664-919), subcloned into the vector pET15b with an N-terminal His₆-tag, was a kind gift from A. Visser (Organon, now Merck). The hAR LBD (residues 664-919), subcloned into the vector pGEX-KG with an N-terminal GST tag, was provided by P. Donner (Bayer-Schering Pharma AG). The hER α LBD (residues 302-553) with an N-terminal His₆-tag was subcloned into pET15b and was provided by A. Visser (Organon, now Merck). The construct pET15-hER β LBD expressing hER β LBD (residues 260-502) with an N-terminal His₆-tag was constructed by subcloning an *Nde*I/*Bam*HI fragment of pET15-hER β , provided by P. Donner (Bayer-Schering-Pharma AG), into pET15b (Novagen). Used primers: 5'-TTTTTTCATATGCTGGACGCCCTGAGCCCCGAGCAG3' and 5'-TTTTTTGGATCCTCACCCGGAAGCACGTGGGCATTCAGCATCTC3'. The 0.24 kb fragment was digested with *Nde*I/*Bam*HI and ligated with the newly double digested pET15b expression vector. The DNA expressing encoding for ER β ₂₆₁₋₅₀₀ was amplified from pET15-hER β by PCR using the forward primer 5'-GAACCATGGACGACGCCCTGAGCCCCGAGCAGCTAGTG-3' and the reverse primer 5'-GGACTCGAG-TTAGTCGTC AAGCACGTGGGCATTCAGCATCTC-3' to introduce the restriction site for *Nco*I and *Xho*I. The primers encode for three extra asp codons, one before the codon for D₂₆₁ and two after L₅₀₀. The expressed LBD thus featured the following sequence: MD[D₂₆₁-L₅₀₀]DD. The PCR fragment was inserted into the *E. coli* expression vector pET16b (Novagen), double digested with the endonucleases *Nco*I and *Xho*I.

His-hAR LBD Expression and Purification. The plasmid was transformed into high-density *Escherichia coli* BL21 (DE3) cells and grown in 6 L TB medium using selection with ampicillin (100 μ g/ml). The cultures were incubated at 37 °C till an OD₆₀₀ of ~1.2 and after cooling down to 18 °C protein expression was induced by adding isopropyl β -D-1-thiogalactopyranoside (IPTG) to a final concentration of 100 μ M. The cells were grown in the presence of 10 μ M DHT (5 α -Androstan-17 β -ol-3-one, Fluka) for 18-20 h at 18 °C and harvested by centrifugation (Beckman Coulter, Avanti J26 XP) at 4500 rpm, 20 min and dissolved in a buffer of 50 mM Tris-HCl, pH 7.5, 500 mM NaCl, 0.4 mM pefabloc (Roche), 50 μ M DHT and 10% glycerol (10 ml/g of cells) or stored at -80 °C till further use. Subsequently, the bacterial suspension was lysed by up to four cycles under a pressure of 80 bar via shear forces using a micro fluidizer (Microfluidizer 1109) and centrifuged (Beckman Coulter, Avanti J25; 20.000 rpm, 30 min, 4 °C). The soluble cell lysate was immobilized on an equilibrated Nickel-NTA agarose column (HisTrap HP, 5 ml, Amersham Biosciences) using Fast Protein Liquid chromatography (Äkta FPLC, Amersham Biosciences) and washed with buffer (50 mM Tris-HCl, pH 7.5, 500 mM NaCl, 10% glycerol, 10 μ M DHT and 10 mM β -Mercaptoethanol) to remove all non-bound proteins. The His₆-AR LBD fusion protein was finally eluted via an imidazole gradient from 0-500 mM imidazol. Fractions containing the fusion protein were combined and desalted on a Sephadex G25 PD-10 column (Amersham

Biosciences) using desalting buffer (50 mM Tris-HCl pH 7.5, 25 mM NaCl, 10 % glycerol and 0.05% β -Octylglycosid). Purity and characterization of the eluted fractions was established by SDS-PAGE using a molecular weight marker (PageRuler PlusPreStained Protein Ladder, Fermentas) and photometric determination of protein concentration using Nanodrop at a wavelength of 280 nm.

GST-hAR LBD Expression and Purification. Expression of the GST-AR LBD was performed as described above for the His-AR LBD. After expression, the cells were lysed with a microfluidizer (4 passes at 600 kPa) in buffer H (50 mM HEPES, pH 7.3, 300 mM NaCl, 5 mM EDTA, 10% glycerol, 100 μ M DHT, 100 μ M PMSF, 10 mM DTT) and centrifuged at 20.000 rpm for 30 min. The soluble cell lysate was immobilized on a glutathione Sepharose 4 Fast Flow affinity matrix (Amersham Biosciences), washed with buffer A (50 mM HEPES, pH 7.3, 300 mM NaCl, 5 mM EDTA, 10% glycerol, 10 μ M DHT, 1 mM DTT) and eluted with buffer A containing 15 mM glutathione. Fractions containing the fusion protein were combined and desalted on a Sephadex G25 PD-10 column (Amersham Biosciences) pre-equilibrated with buffer A. Purity and characterization of the eluted fractions was established by SDS-PAGE using a molecular weight marker (PageRuler PlusPreStained Protein Ladder, Fermentas) and photometric determination of protein concentration using Nanodrop at a wavelength of 280 nm. In case of cleavage of the GST moiety with thrombin over night at 4 °C a final purification step was followed using a HiTrap SP cation exchange column (Amersham Biosciences).

hER LBD Expression and Purification. The resulting plasmids (pET15-His-hER α LBD and pET15b-His-hER β LBD) were each transformed into *E. coli* BL21 (DE3) cells, and grown in 4 L TB medium using selection with ampicillin (100 μ g/ml). The cultures were incubated at 37 °C till an OD₆₀₀ of ~1.2 and after cooling down to 15 °C, protein expression was induced by adding IPTG to a final concentration of 100 μ M and the cells were grown in the presence of 10 μ M E₂ (β -Estradiol, Serva) for 18-20 h at 15 °C and harvested by centrifugation (4500 rpm, 20 min, 4 °C) and the pellet was stored at -80 °C till further use. Cells were lysed with a micro fluidizer (4 passes at 600 kPa) in lysis buffer (PBS containing 26.8 mM KCl, 14.7 mM KH₂PO₄, 78.1 mM Na₂HPO₄, 370 mM NaCl, 40 mM Imidazol, pH 8 and 10 % glycerol) and centrifuged (20.000 rpm, 30 min, 4 °C). The soluble cell lysate was immobilized on an equilibrated Nickel-NTA agarose column (HisTrap HP, 5 ml, Amersham Biosciences), washed with lysis buffer and eluted via a gradient using elution buffer (PBS containing 26.8 mM KCl, 14.7 mM KH₂PO₄, 78.1 mM Na₂HPO₄, 370 mM NaCl, 500 mM Imidazol pH 8, 10% glycerol). Fractions containing the fusion protein were combined and desalted on a Sephadex G25 PD-10 column (Amersham Biosciences) using desalting buffer (20 mM Tris, 25 mM NaCl, 10 % glycerol and 0.05 % β -Octylglycosid). Determination of purification efficiency was performed by SDS gel electrophoresis. The concentration of both proteins was quantified using Nanodrop at a wavelength of 280 nm.

Expression and Purification of the ER β LBD for crystallization. The ER β LBD₂₆₁₋₅₀₀ was overexpressed from high-density culture of *E. coli* BL21 (DE3) host cells (Statagene). Harvested cells were lysed using a micro fluidizer in a buffer containing 20 mM Tris-HCl pH 7.5, 0.5 M NaCl, 5 mM DTT, and 1 mM EDTA (10 ml/g of cells). Clarified lysate was flowed through a pre-equilibrated 3 ml estradiol-Sepharose column (PTI Research, Inc.) and washed with 250 ml of 10 mM Tris-HCl, pH 7.5 containing 0.5 M NaCl and 1 mM EDTA (buffer A). The column was then re-equilibrated with 50 ml of 10 mM Tris-HCl, pH 7.5, 0.2 M NaCl, and 1 mM EDTA (buffer B), and then the protein was carboxymethylated using 50 ml of buffer B containing 5 mM iodoacetic

acid. After incubation for 1 h at 4 °C, the column was washed by 400 ml of buffer A, followed by elution in 20 ml buffer A containing 50-200 µM estradiol. Subsequently partial contaminations were removed by size exclusion chromatography (Sephadex 75, HiLoad 26/60, GE Healthcare Biosciences). Fractions containing the purified ERβ LBD₂₆₁₋₅₀₀ were recombined and the buffer was exchanged into buffer B containing 5 mM DTT by passing the solution through a Sephadex G25 PD-10 column (Amersham Biosciences). Finally, the eluted ERβ LBD₂₆₁₋₅₀₀ was concentrated to 10-12 mg/ml using Amicon ultra centrifuge tubes (MWCO 10 kDa) and characterized by SDS-PAGE and photometric determination of protein concentration using Nanodrop at a wavelength of 280 nm.

Cleavage of the affinity-tag. If necessary enzymatic cleavage of the His₆-tag or GST-tag was performed using a thrombin protease (1 unit/µl; Amersham Biosciences). One cleavage unit was incubated with 100 µg of fusion protein in desalting buffer at room temperature over night. After cleavage the thrombin was removed from the sample by addition of benzamidine sepharose beads (GE Healthcare), which bind thrombin. The beads were three times washed in distilled water before usage. The beads (1 µl per 1 µl protein) were incubated with the sample for about 30 minutes at 4 °C under slow stirring followed by a 10 minutes centrifugation step at 4500 rpm to remove the sepharose.

Fluorescence Polarization Assay hAR LBD. Purified hAR LBD was serially diluted into assay buffer (20 mM HEPES, pH 7.2, 50 mM KCl, 1 mM EDTA, 10 µM DHT, 1 mM DTT, 1 mg/ml BSA) with final concentrations of 10 µM DHT and 0.1 µM fluorescein-labeled peptide FI-CSSRFESLFAAGEKESR^[30] using a black 384-well plate (Perkin Elmer, Optiplate-384 F). The final volume in each well was 50 µl. Each plate also included the following controls: the empty well, 0.1 µM fluorescent peptide in reaction buffer and reaction buffer containing the same final concentration of DMSO used to solve the ligand. The samples were mixed by pipetting up and down and subsequently the plate was centrifuged (5 min, 7000 rpm, 4 °C) to remove bubbles. The reaction was incubated at 4 °C for 1 h in the dark. The fluorescence polarization was measured at 23 °C using a plate reader (Safire², Tecan) with an excitation at 470 nm, an emission at 519 nm, a gain of 80 and 50 reads per well. The G-factor (1.154) was obtained by collecting parallel and perpendicular components of fluorescence from an 8 nM solution of fluorescein in water. The data derived from the fluorescence polarization (in millipolarization, mP) were normalized, plotted against increasing concentrations of the hAR LBD and then fitted with ORIGIN 7 (Scientific Graphing and Analysis Software, OriginLab Corp.) using nonlinear regression analysis with a sigmoidal dose-response (variable slope) equation to determine the K_D value of the fluorescein labeled peptide – hAR LBD complex. With these experimental data and the use of equation 2 a K_D of 0.96 µM was determined for the fluorescent reference peptide – AR LBD interaction.

$$P = P_0 + \frac{(P_{fm} - P_0)(A_0 + B_0 + K_d - \sqrt{(A_0 + B_0 + K_d)^2 - 4A_0B_0})}{2A_0} \quad (\text{Eq. 2})$$

where P is the measured polarization value, P_0 is the polarization of the free fluorescent ligand, P_{fm} is the polarization of the bound ligand, A_0 is the total concentration of fluorescent peptide, B_0 is the protein concentration and K_D is the dissociation constant of the protein-peptide complex. Each data point represents the average of an experimental condition performed in at least triplicate. Z' factor analysis was performed as

recently described by Zhang *et al*^[66]. Assays with a Z' factor between 0.5 and 1.0 are considered to be reliable and robust.

Fluorescence Polarization Assay hER LBD. Purified hER LBD was serially diluted into assay buffer (20 mM HEPES, pH 7.3, 100 mM NaCl, 1 mM EDTA, 10 μ M E₂, 1 mM DTT, 0.005 % Nonident P40) with final concentrations of 10 μ M E₂ and 0.1 μ M fluorescein-labeled peptide FI-CLTERHKILHRLLQEGSPSD (from SRC1 Box 2) in a black 384-well plate (Perkin Elmer, Optiplate-384 F). The final volume in each well was 50 μ l. The measurement of the polarization and the calculation of the K_D values were performed as described for the AR LBD. With the use of equation 1 a K_D of 0.77 μ M was determined for the fluorescent reference peptide in case of the ER α LBD and a K_D of 0.297 μ M was measured for the ER β LBD.

Fluorescence Polarization Competitive Displacement Assay hAR LBD. The optimized assay mixture contained 0.1 μ M fluorescein-labeled peptide FI-CSSRFESLFAGEKESR^[30] and 1 μ M of purified hAR LBD in reaction buffer (20 mM HEPES, pH 7.2, 50 mM KCl, 1 mM EDTA, 10 μ M DHT, 1 mM DTT, 1 mg/ml BSA). Inhibition experiments were performed in 384-well plates (Perkin Elmer, Optiplate-384 F) by adding 40 μ l of the reaction mixture to 10 μ l of increasing amounts of CBI (diluted in reaction buffer). Samples without inhibitor as well as samples containing only FI-labeled peptide in reaction buffer were used as controls. The reaction was incubated at room temperature for 1 h in the dark (measurements repeated after 24 h and stored at 4 °C showed no significant difference in the IC₅₀ values). The fluorescence polarization was measured at 23 °C using a plate reader (Safire², Tecan) with an excitation at 470 nm and an emission at 519 nm. The polarization data (in millipolarization, mP) were normalized, plotted against the log₁₀ of increasing concentrations of the inhibitor and then fitted with a Klotz binding model^[67] to a sigmoid curve using ORIGIN 7.5 (Scientific Graphing and Analysis Software, OriginLab Corp.) to determine the K_I and IC₅₀ value of the inhibitor. The competitive binding of the CBIs were measured at least two times in independent duplo experiments. The fluorescence polarization data of each CBI was used to calculate the IC₅₀ using equation 3.

$$P = P_{\min} + \frac{(P_{\max} - P_{\min})}{\left(1 + 10^{(\log(x) - \log(IC_{50}))}\right)} \quad (\text{Eq. 3})$$

where P is the polarization, P_{\min} is the minimum value of polarization and P_{\max} is the maximum value of polarization. With the aid of equation 4, the K_D of the fluorescein labeled peptide and the IC₅₀ values of each miniprotein, the K_I values for the different miniproteins could be determined^[68].

$$K_i = \frac{IC_{50}}{1 + \frac{A_0 \cdot (y_0 + 2)}{2 \cdot K_d \cdot (y_0 + 1)} + y_0} - K_d \cdot \left(\frac{y_0}{y_0 + 2} \right) \quad (\text{Eq. 4})$$

where IC₅₀ is the half maximal inhibitor concentration, K_D is the dissociation constant of the protein - labelled peptide complex and y₀ is the initial bound-to-free concentration ratio for labelled peptide.

$$y_0 = \frac{AB}{A} = \frac{B}{Kd}$$

where AB is the concentration of protein-peptide complex and A the concentration of unbound peptide.

Fluorescence Polarization Competitive Displacement Assay hER LBD. The optimized reaction mixture contained 0.1 μM fluorescein-labeled peptide FL-CLTERHKILHRLLQEGSPSD (from SRC1 Box 2) and 1 μM of purified hER LBD in reaction buffer (20 mM HEPES, pH 7.2, 50 mM KCl, 1 mM EDTA, 10 μM E_2 , 1 mM DTT, 1 mg/ml BSA). Inhibition experiments were performed and evaluated as described above for the AR LBD.

Cofactor recruitment FRET assay. The optimized reaction mixture contained 10 nM of purified His-ER α -LBD and ER β -LBD respectively, 1.25 nM europium (Eu) α -His antibody (Perkin Elmers), 100 nM of biotin labeled SRC1 α -4 peptide and 80 nM streptavidin-APC (Perkin Elmers) in the reaction buffer (50 mM Tris-HCl, pH 7.4, 50 mM KCl, 1 mM EDTA, 0.1 mg/ml BSA, 1 mM DTT). Each experiment was performed in 384-well plates (384 Packard Optiplate) by adding 20 μl of the receptor plus Eu labeled α -His antibody to 10 μl of peptide plus streptavidin-APC. Dilutions of E_2 ranging from 10 μM to 0.01 nM were added to the solution. The reactions were mixed and routinely incubated overnight at 4 $^\circ\text{C}$. The samples were counted using an EnVision Counter with a setting of 60 μsec time delay, excitation at 320 nm, and emission at 665 nm. Data were collected from counts at 665 nm and directly used for sigmoid curve plotting using Origin 7.5 (Scientific Graphing and Analysis Software, OriginLab Corp.). Each time point is the average of two independent measurements.

Isothermal Titration Calorimetry. ITC experiments were performed using a VP-ITC microcalorimeter (Microcal, Inc.). Purified ER α/β LBD was dialyzed extensively against thoroughly degassed phosphate buffered saline (PBS) containing 1.37 M NaCl, 26.8 mM KCl, 14.7 mM KH_2PO_4 , 78.1 mM Na_2HPO_4 , pH 7.5 and 100 μM E_2 . Peptides were dissolved in the same buffer also containing 100 μM of E_2 . Titrations were performed at 26 $^\circ\text{C}$ and consisted of 35 injections of 8 μl peptide for 20 s and were separated by 240 s. Each titration contained 20 μM ER α/β LBD in the sample cell. Concentrations of SRC-1 box 2 and Apa-3 peptides in the injection syringe ranged from 1 μM to 200 μM . The initial injection was discarded from the data sheet in order to remove the effect of titrant diffusion across the syringe tip during the equilibration process. The experimental data were fitted to a one site binding model using Origin 5.0 (Microcal, Inc.) to determine the binding constant (K_A), apparent stoichiometry (N), and changes in binding free energy (ΔG), enthalpy (ΔH), and entropy (ΔS). For all experiments, the quantity, $c = K_A \text{Mt}(0)$, where $\text{Mt}(0)$ is the initial macromolecular concentration, was observed to be in the range of $14 < c < 134$. Control experiments were performed by making the identical injections of peptides into a cell containing buffer with no protein.

Co-crystallization of the ligand-bound NR LBD in complex with miniprotein. For crystallization, the AR LBD/DHT complex was concentrated to 4 mg/ml in a buffer containing 50 mM Tris-HCl pH 7.5, 0.5 mM NaCl, 10% glycerol, 1 mM EDTA, 1 mM DTT, 0.05% NaAzide and 10 μM DHT and then mixed with peptide at a molecular ratio of 2:1, 3:1, 4:1, 5:1 peptide to protein – ligand complex. Prior, the peptide was dissolved in crystallization buffer at a concentration of 10 mM and extensively dialyzed against the same buffer for three days at 4 $^\circ\text{C}$ using a dialysis membrane from Spectra/Por (MWCO: 1 kDa). In order to crystallize the AR LBD

complex, initial screenings employing JCSG+, JCSG Core I, JCSG Core II, JCSG Core III and JCSG Core IV from Qiagen were performed at 20 and 4 °C using the sitting drop vapor diffusion. The protein concentration in the setups was 4 mg/ml; 0.1 µl protein solutions were automatically mixed with 0.1 µl reservoir solution in 96 well Corning pZero plates using a phoenix pipetting robot. The sitting drops were equilibrated against reservoirs with a volume of 70 µl. After one to two days, rod shaped and cylindrical crystals appeared in several conditions typically containing 20% PEG 3350 and reaching a maximal size of 90 x 13 x 10 µm within three days. similar screening experiments were carried out with the ERβ LBD/E₂/LXXLL peptide complex. The receptor was concentrated 10 12 mg/ml in a buffer containing 0.2 M NaCl, 1 mM EDTA, 5 mM DTT and 10 mM Tris-HCl at pH 7.5 and then mixed with the SRC-1 Box 2 peptide at a molecular ratio of 2:1 and 5:1 peptide to protein – ligand complex. Prior, the peptide was dissolved in crystallization buffer to a concentration of 8.97 mM and extensively dialyzed against the same buffer for three days at 4 °C using a dialysis membrane from Spectra/Por (MWCO: 1 kDa). The same screening conditions as for the AR LBD were used for crystallization, leading to crystals of similar size and shape growing in 0.1 Bicine pH 9 and 20% PEG 6000. These crystals reached a final size of 190 x 20 x 20 µm within five days. Reproduction of these crystals was performed in 24 well EasyXtal DG-tools (Qiagen) using the hanging drop vapor diffusion method. Drops with a volume of 2 µl using a variable reservoir to protein ratio were manually mixed at 4 °C and equilibrated against reservoirs (Qiagen refill hits) with a volume of 1 ml.

3.9 References

- [1] M. J. Tsai, B. W. O'Malley, *Annu Rev Biochem* **1994**, *63*, 451.
- [2] H. Gronemeyer, J. A. Gustafsson, V. Laudet, *Nature Reviews Drug Discovery* **2004**, *3*, 950.
- [3] W. Bourguet, P. Germain, H. Gronemeyer, *Trends Pharmacol Sci* **2000**, *21*, 381.
- [4] L. Nagy, J. W. Schwabe, *Trends Biochem Sci* **2004**, *29*, 317.
- [5] K. W. Nettles, G. L. Greene, *Annu Rev Physiol* **2005**, *67*, 309.
- [6] D. M. Heery, E. Kalkhoven, S. Hoare, M. G. Parker, *Nature* **1997**, *387*, 733.
- [7] B. He, J. A. Kempainen, E. M. Wilson, *Journal of Biological Chemistry* **2000**, *275*, 22986.
- [8] A. L. Rodriguez, A. Tamrazi, M. L. Collins, J. A. Katzenellenbogen, *J Med Chem* **2004**, *47*, 600.
- [9] C. Chang, J. D. Norris, H. Gron, L. A. Paige, P. T. Hamilton, D. J. Kenan, D. Fowlkes, D. P. McDonnell, *Mol Cell Biol* **1999**, *19*, 8226.
- [10] J. M. Hall, C. Y. Chang, D. P. McDonnell, *Molecular Endocrinology* **2000**, *14*, 2010.
- [11] T. R. Geistlinger, R. K. Guy, *J Am Chem Soc* **2003**, *125*, 6852.
- [12] A. M. Leduc, J. O. Trent, J. L. Wittliff, K. S. Bramlett, S. L. Briggs, N. Y. Chirgadze, Y. Wang, T. P. Burris, A. F. Spatola, *Proc Natl Acad Sci U S A* **2003**, *100*, 11273.
- [13] A. K. Galande, K. S. Bramlett, J. O. Trent, T. P. Burris, J. L. Wittliff, A. F. Spatola, *Chembiochem* **2005**, *6*, 1991.
- [14] M. Carraz, W. Zwart, T. Phan, R. Michalides, L. Brunsveld, *Chem Biol* **2009**, *16*, 702.
- [15] N. B. Mettu, T. B. Stanley, M. A. Dwyer, M. S. Jansen, J. E. Allen, J. M. Hall, D. P. McDonnell, *Mol Endocrinol* **2007**, *21*, 2361.
- [16] D. Shao, T. J. Berrodin, E. Manas, D. Hauze, R. Powers, A. Bapat, D. Gonder, R. C. Winneker, D. E. Frail, *J Steroid Biochem Mol Biol* **2004**, *88*, 351.
- [17] L. A. Arnold, E. Estebanez-Perpina, M. Togashi, N. Jouravel, A. Shelat, A. C. McReynolds, E. Mar, P. Nguyen, J. D. Baxter, R. J. Fletterick, P. Webb, R. K. Guy, *Journal of Biological Chemistry* **2005**, *280*, 43048.
- [18] J. Becerril, A. D. Hamilton, *Angew Chem Int Ed Engl* **2007**, *46*, 4471.
- [19] J. R. Gunther, T. W. Moore, M. L. Collins, J. A. Katzenellenbogen, *ACS Chem Biol* **2008**, *3*, 282.
- [20] A. L. LaFrate, J. R. Gunther, K. E. Carlson, J. A. Katzenellenbogen, *Bioorg Med Chem* **2008**, *16*, 10075.
- [21] A. A. Parent, J. R. Gunther, J. A. Katzenellenbogen, *J Med Chem* **2008**, *51*, 6512.
- [22] A. B. Williams, P. T. Weiser, R. N. Hanson, J. R. Gunther, J. A. Katzenellenbogen, *Org Lett* **2009**, *11*, 5370.
- [23] T. W. Moore, C. G. Mayne, J. A. Katzenellenbogen, *Mol Endocrinol* **2009**.

- [24] J. Y. Hwang, L. A. Arnold, F. Zhu, A. Kosinski, T. J. Mangano, V. Setola, B. L. Roth, R. K. Guy, *J Med Chem* **2009**, *52*, 3892.
- [25] J. R. Gunther, Y. Du, E. Rhoden, I. Lewis, B. Revenaugh, T. W. Moore, S. H. Kim, R. Dingledine, H. Fu, J. A. Katzenellenbogen, *J Biomol Screen* **2009**, *14*, 181.
- [26] C. Y. Chang, J. Abdo, T. Hartney, D. P. McDonnell, *Molecular Endocrinology* **2005**, *19*, 2478.
- [27] E. Estebanez-Perpina, L. A. Arnold, P. Nguyen, E. D. Rodrigues, E. Mar, R. Bateman, P. Pallai, K. M. Shokat, J. D. Baxter, R. K. Guy, P. Webb, R. J. Fletterick, *Proc Natl Acad Sci U S A* **2007**, *104*, 16074.
- [28] J. R. Gunther, A. A. Parent, J. A. Katzenellenbogen, *ACS Chem Biol* **2009**, *4*, 435.
- [29] B. He, R. T. Gampe, Jr., A. J. Kole, A. T. Hnat, T. B. Stanley, G. An, E. L. Stewart, R. I. Kalman, J. T. Minges, E. M. Wilson, *Mol Cell* **2004**, *16*, 425.
- [30] E. Hur, S. J. Pfaff, E. S. Payne, H. Gron, B. M. Buehrer, R. J. Fletterick, *PLoS Biol* **2004**, *2*, E274.
- [31] J. W. Neidigh, R. M. Fesinmeyer, N. H. Andersen, *Nature Structural Biology* **2002**, *9*, 425.
- [32] M. G. Woll, E. B. Hadley, S. Mecozzi, S. H. Gellman, *J Am Chem Soc* **2006**, *128*, 15932.
- [33] A. M. Hodges, A. Schepartz, *J Am Chem Soc* **2007**, *129*, 11024.
- [34] F. Bernal, A. F. Tyler, S. J. Korsmeyer, L. D. Walensky, G. L. Verdine, *J Am Chem Soc* **2007**, *129*, 2456.
- [35] C. Vita, J. Vizzavona, E. Drakopoulou, S. Zinn-Justin, B. Gilquin, A. Menez, *Biopolymers* **1998**, *47*, 93.
- [36] H. Kolmar, *FEBS J* **2008**, *275*, 2684.
- [37] F. Stricher, L. Martin, C. Vita, *Methods Mol Biol* **2006**, *340*, 113.
- [38] H. Kolmar, *Curr Opin Pharmacol* **2009**, *9*, 608.
- [39] C. P. Sommerhoff, O. Avrutina, H. U. Schmoldt, D. Gabrijelcic-Geiger, U. Diederichsen, H. Kolmar, *J Mol Biol* **2010**, *395*, 167.
- [40] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, *Nucleic Acids Res* **2000**, *28*, 235.
- [41] J. H. Pease, D. E. Wemmer, *Biochemistry* **1988**, *27*, 8491.
- [42] K. N. Srinivasan, V. Sivaraja, I. Huys, T. Sasaki, B. Cheng, T. K. Kumar, K. Sato, J. Tytgat, C. Yu, B. C. San, S. Ranganathan, H. J. Bowie, R. M. Kini, P. Gopalakrishnakone, *Journal of Biological Chemistry* **2002**, *277*, 30040.
- [43] C. Vita, E. Drakopoulou, J. Vizzavona, S. Rochette, L. Martin, A. Menez, C. Roumestand, Y. S. Yang, L. Ylisastigui, A. Benjouad, J. C. Gluckman, *Proc Natl Acad Sci U S A* **1999**, *96*, 13091.
- [44] B. Chagot, C. Pimentel, L. Dai, J. Pil, J. Tytgat, T. Nakajima, G. Corzo, H. Darbon, G. Ferrat, *Biochem J* **2005**, *388*, 263.
- [45] R. B. Merrifield, *Journal of the American Chemical Society* **1963**, *85*, 2149.
- [46] L. A. Carpino, G. Y. Han, *Journal of the American Chemical Society* **1970**, *92*, 5748.
- [47] J. Hachmann, M. Lebl, *Biopolymers* **2006**, *84*, 340.
- [48] I. Coin, P. Schmieder, M. Bienert, M. Beyermann, *Biopolymers* **2007**, *88*, 565.
- [49] V. Cavallaro, P. E. Thompson, M. T. W. Hearn, *Journal of Peptide Science* **2001**, *7*, 529.
- [50] F. D. Sonnichsen, J. E. Van Eyk, R. S. Hodges, B. D. Sykes, *Biochemistry* **1992**, *31*, 8790.
- [51] A. J. Nicoll, C. J. Weston, C. Cureton, C. Ludwig, F. Dancea, N. Spencer, O. S. Smart, U. L. Gunther, R. K. Allemann, *Org Biomol Chem* **2005**, *3*, 4310.
- [52] Greenfield, N. G. D. Fasman, *Biochemistry* **1969**, *8*, 4108.
- [53] J. P. Hennessey, W. C. Johnson, *Biochemistry* **1981**, *20*, 1085.
- [54] S. W. Provencher, J. Glockner, *Biochemistry* **1981**, *20*, 33.
- [55] R. W. Woody, *Methods Enzymol* **1995**, *246*, 34.
- [56] C. J. Weston, C. H. Cureton, M. J. Calvert, O. S. Smart, R. K. Allemann, *Chembiochem* **2004**, *5*, 1075.
- [57] B. Vaz, S. Mocklinghoff, S. Folkertsma, S. Lusher, J. de Vlieg, L. Brunsveld, *Chem Commun (Camb)* **2009**, 5377.
- [58] R. N. Armstrong, *Chem Res Toxicol* **1991**, *4*, 131.
- [59] D. B. Smith, K. S. Johnson, *Gene* **1988**, *67*, 31.
- [60] E. R. LaVallie, J. M. McCoy, *Curr Opin Biotechnol* **1995**, *6*, 501.
- [61] U. K. Laemmli, *Nature* **1970**, *227*, 680.
- [62] W. J. Checovich, R. E. Bolger, T. Burke, *Nature* **1995**, *375*, 254.
- [63] D. M. Jameson, W. H. Sawyer, *Biochemical Spectroscopy* **1995**, *246*, 283.
- [64] T. Heyduk, Y. Ma, H. Tang, R. H. Ebright, *Methods Enzymol* **1996**, *274*, 492.
- [65] S. Y. Tetin, T. L. Hazlett, *Methods* **2000**, *20*, 341.
- [66] J. H. Zhang, T. D. Y. Chung, K. R. Oldenburg, *Journal of Biomolecular Screening* **1999**, *4*, 67.
- [67] I. M. Klotz, J. M. Longfellow, O. H. Johnson, *Science* **1946**, *104*, 264.
- [68] T. J. Burke, K. R. Loniello, J. A. Beebe, K. M. Ervin, *Comb Chem High Throughput Screen* **2003**, *6*, 183.
- [69] S. Pegoraro, S. Fiori, J. Cramer, S. Rudolph-Bohner, L. Moroder, *Protein Sci* **1999**, *8*, 1605.
- [70] W. Kabsch, *Journal of Applied Crystallography* **1993**, *26*, 795.

- [71] K. S. Bramlett, Y. F. Wu, T. P. Burris, *Molecular Endocrinology* **2001**, *15*, 909.
- [72] M. A. Iannone, T. G. Consler, K. H. Pearce, J. B. Stimmel, D. J. Parks, J. G. Gray, *Cytometry* **2001**, *44*, 326.
- [73] J. C. Sheehan, G. P. Hess, *Journal of the American Chemical Society* **1955**, *77*, 1067.
- [74] K. Mullis, F. Faloona, S. Scharf, R. Saiki, G. Horn, H. Erlich, *Cold Spring Harb Symp Quant Biol* **1986**, *51 Pt 1*, 263.

Chapter 4

Protein Semi-Synthesis and Evaluation of the Phosphorylated Estrogen Receptor β Ligand Binding Domain

Abstract: There is a need for NR constructs featuring post-translational modifications (PTMs) for *in vitro* studies, not only to enable a better comparison with *in vivo* data, but also to allow a detailed investigation how NR functioning is regulated via PTMs. Using the Estrogen Receptor (ER β) as an example, in this chapter Expressed Protein Ligation (EPL) is reported as methodology to generate a phosphorylated LBD and to demonstrate the influence of a certain phosphorylated amino acid on cofactor binding.

4.1 Introduction

Although nuclear receptors (NRs) are ligand-dependent multidomain transcription factors, their expression and many of their activities are highly regulated by post-translational modifications (PTMs), including phosphorylation^[1]. While the process of ligand binding on the subsequent steps of NR action is relatively well understood on the molecular level^[2], the influence of NR PTMs on ligand binding and on the formation of multiple protein complexes is, however, largely unknown. In particular, a molecular understanding of how the plethora of NR PTMs regulates the crosstalk between NRs and cofactor proteins, resulting in epigenetic chromatin modifications, is missing. Typically, biochemical and structural studies are performed on the isolated NR Ligand Binding Domain (LBD)^[3], which is obtained via expression in bacteria and does not feature PTMs^[4]. Novel approaches are therefore required to obtain a NR LBD, post-translationally modified at specific positions, for *in vitro* studies.

Expressed Protein Ligation (EPL)^[5] could present a synthetic entry into NR constructs featuring specifically introduced PTMs, and thus enabling the study of the crosstalk between NRs and cofactor proteins, as well as other phenomena, on the molecular level. EPL, elaborating on the native chemical ligation (NCL)^[6] strategy, has become the most popular tool for site-specific modification of proteins^[5a, 7]. NCL is a chemoselective reaction occurring spontaneously between a peptide α -thioester (N-terminal segment) and an N-terminal cysteine containing peptide (C-terminal segment), to form a native peptide bond. As an extension of the NCL strategy, EPL enables the generation of a protein α -thioester through the development of expression systems based on self-cleavable intein domains. Combining chemical peptide synthesis with protein expression strategies^[7a] has permitted ready generation of semi-synthetic proteins that site-specifically incorporate unnatural amino acids or side-chain modifications (Figure 1).

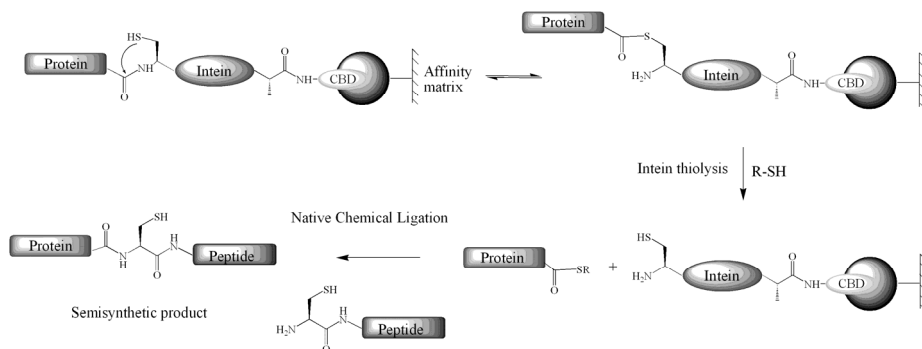


Figure 1: Schematic representation of the generation of recombinant protein α -thioesters followed by expressed protein ligation (EPL).

A PTM on NRs that represents a suitable and interesting target to investigate is the ER β tyrosine 488 (Y₄₈₈), which in analogy to ER α Y₅₃₇^[8], is regulated via phosphorylation^[8d, 9]. The cross-talk of the estrogen receptors with c-Src kinase via phosphorylation of this tyrosine has been intensively investigated in cellular studies. The phosphorylation is, for example, critical for the assembly of the ER β /AR/Src complex in the cytosol as it activates the Src-dependent pathway^[10]. Although the role of the tyrosine phosphorylation in inducing interactions with SH2 domains in the cytosol is relatively well established via cellular studies, its role in protein-protein interactions in the nucleus with epigenetic enzymes is much less clear. Cellular studies have indicated that phosphorylation is critical for DNA binding and receptor dimerization^[8b], but other studies have observed only fine-tuning effects from the presence of a phosphorylatable tyrosine^[11]. Thus, the exact function of a phosphorylation of the ER β at tyrosine 488 remains controversial, supporting the need for new molecular insights in how tyrosine phosphorylation influences the ER β function. Furthermore, the participation of phosphorylation at this site on coactivator binding has not yet been investigated in detail. Nevertheless, a regulating effect of ER β Y₄₈₈ phosphorylation on interactions with cofactor proteins is likely since this tyrosine is located in close proximity to the cofactor interacting surface of the LBD^[12], activating function 2 (AF-2) (Figure 2). Biochemical and structural studies on the specific phosphorylated ER β LBD promise to be a valuable tool to give new important molecular insights in how phosphorylation of the ER β LBD on tyrosine 488 regulates ER β function.

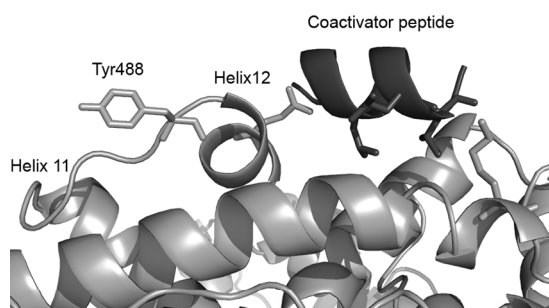


Figure 2: Zoom-in on the co-crystal structure of the ER β LBD with the SRC1 box 2 coactivator peptide (black) and estradiol (solved in our lab, deposition pending). Tyrosine 488 is located in the loop between helix 11 and helix 12 in close proximity to the cofactor interacting surface.

In this chapter, the protein semi-synthesis of the ER β LBD and its selective and homogeneous phosphorylation on tyrosine 488 via expressed protein ligation (EPL) is reported. By successfully solving the crystal structure of the semi-synthetic phosphorylated

ER β LBD as well as by several biochemical studies a demonstration could be given of how this phosphorylation influences the interaction of the ER β LBD with other proteins.

4.2 Protein Semi-Synthesis of the Phosphorylated ER β LBD

The classical way to achieve PTM of a protein target *in vitro* relies on the application of specific enzymes to the protein target. C-Src has been assumed to be the kinase regulating the phosphorylation of ER α Y₅₃₇ and ER β Y₄₈₈^[8b, 13]. In cooperation with Arie Visser and Elisabeth van der Vaart (Organon, now Merck), biochemical efforts were undertaken to phosphorylate both proteins using commercially available c-Src kinase. Incubation of the enzyme with the ER α LBD resulted in protein phosphorylation (Figure 3). However, the yield and selectivity of this reaction and thus the homogeneity of the sample could not be determined. Phosphorylation of the ER β LBD via the c-Src kinase on the other hand could not be detected at all.

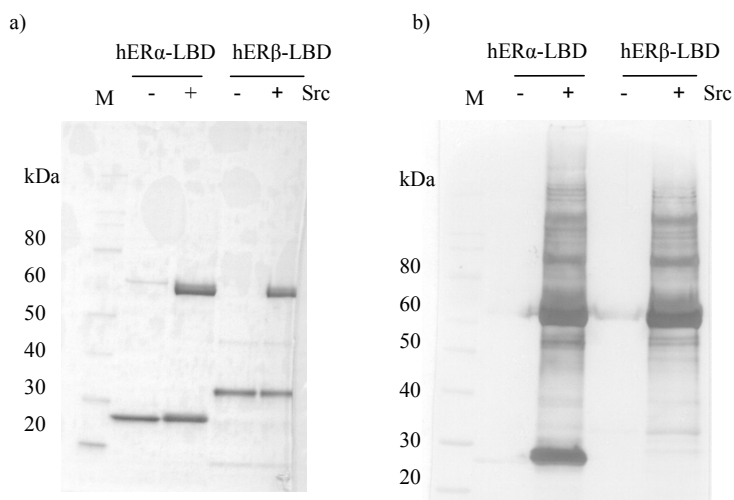


Figure 3: (a) SDS-PAGE^[14] gel (4-10%; Coomassie stained) of the ER α (27 kDa) and ER β LBD (31 kDa) in the absence and in the presence of the Src enzyme (61.7 kDa); (b) Western-Blot (Nitrocellulose; DAB stained) using an α -phosphotyrosine antibody of the ER α and ER β LBD without and in the presence of the Src enzyme. No band could be detected for the phosphorylation of the ER β LBD.

These results clarified that an enzymatic phosphorylation of tyrosine 488 is not trivial and supported the decision to generate the homogeneously and selectively Y₄₈₈ phosphorylated ER β LBD via EPL. This semi-synthetic approach would avoid possible

partial or unselective phosphorylations, occurring when (other) kinases are applied or when the construct is expressed in eukaryotic systems.

In order to use EPL as a method to generate the phosphorylated ER β LBD on tyrosine 488 the protein has to be separated into two segments, one containing a C-terminal thioester and the other containing an N-terminal cysteine. As tyrosine 488 is located at the end of the loop between helices 11 and 12 of the ER β LBD, we chose the naturally occurring cysteine C₄₈₁ at the C-terminus of helix 11 as the optimal ligation junction. Helices 1-11 of the ER β LBD (aa 260-480) were generated as C-terminal thioester excluding the loop and Helix 12 (Figure 4). Helix 12 represents the C-terminal and the most flexible part of the LBD^[15] and it was speculated that removal of this helix would not significantly destabilize the correctly folded state of the expressed protein.

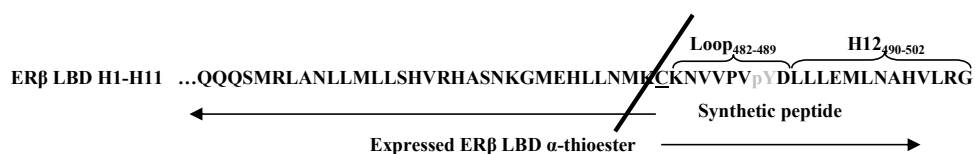


Figure 4: Position of tyrosine 488 in the ER β LBD in close proximity of a cysteine at position 481.

The generation of recombinant proteins bearing C-terminal thioesters is based on protein splicing. Protein splicing is a posttranslational autocatalytic process in which an internal protein domain, a so-called intein, is spliced from a precursor protein in a series of intramolecular rearrangements^[16]. Both the flanking protein domains (exteins) will be ligated at the same time through covalent peptide binding. Engineering of mutated inteins led to the interruption of the splicing reaction at a certain point and enabled the selective splicing of only one extein^[17]. Proteins that are expressed as N-terminal fusions to such a controllable self-splicing intein-affinity domain can be cleaved by thiols via an intermolecular transthioesterification, releasing the recombinant protein containing a C-terminal thioester^[5]. Commercial *Escherichia coli* (*E. coli*) expression and purification systems have been developed that combine intein-catalyzed thiolysis of fusion proteins with an affinity-based purification tag to obtain thioester-terminated recombinant proteins^[18].

In this work, the DNA sequence encoding for the N-terminal part of the ER β LBD (aa 260-480) was subcloned into the expression vector pTWIN1 in-frame with the sequence of an intein-chitin binding domain (CBD)^[19] and expressed in appropriate *E. coli* cells. The resulting ER β LBD fusion protein was immobilized via affinity chromatography on a chitin resin and treated with a thiol containing buffer. The chosen thiol, 2-mercaptoethanesulfonic

acid (MESNA), is small enough to enter into the catalytic pocket of the intein and is able to generate stable α -thioesters *in situ* through transthioesterification^[20] that are highly reactive when exposed to cysteines. A further advantage of MESNA compared to other thiols is its high water solubility, which makes it easy to remove the MESNA again after the reaction. During overnight incubation of the ER β LBD-intein-CBD fusion protein bound to the chitin resin, MESNA induces the cleavage of the thioester bond via connecting the C-terminal residue of the protein to the side chain thiol group of the first residue in the intein.

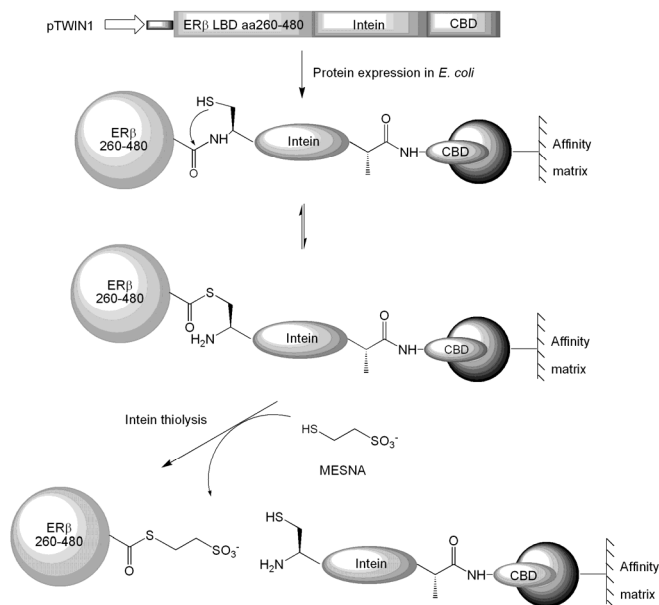


Figure 5: Schematic representation of the purification of the ER β LBD featuring an α -thioester on its C-terminus. The ER β LBD is cleaved from the intein-CBD fusion protein whilst bound to the affinity matrix by the thiol MESNA to give the corresponding protein α -thioester.

The cleaved ER β LBD featuring an α -thioester on its C-terminus was isolated in a yield of approximately 10 mg L⁻¹. SDS-PAGE and ESI analysis^[21] showed the presence of a single protein with a mass of 25185 Da that is consistent with the truncated ER β LBD₂₆₀₋₄₈₀ – MESNA product. Purification of the truncated ER β LBD α -thioester was successful, both for the apo-protein and for the agonist-bound state with estradiol (E₂).

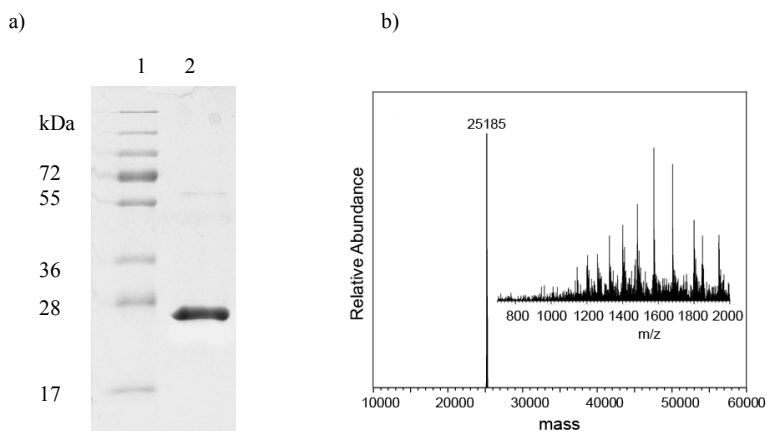


Figure 6: Characterization of the truncated ER β LBD α -thioester; a) SDS-PAGE gel (15 %; Coomassie stained) lane 1: Molecular weight marker; Lane 2: ER β LBD - MESNA ; b) Deconvoluted mass spectrum of ER β LBD₂₆₀₋₄₈₀ - MESNA (calculated mass: 25179 Da), inset: m/z spectrum.

The C-terminal part of the ER β LBD (aa 481-502) including the loop and helix 12 was synthesized via classical solid-phase peptide synthesis (SPPS)^[22], both for the phosphorylated (pY) and, for comparison, for the non-phosphorylated form. The phosphotyrosine residue was incorporated in the commercially available form of Fmoc-Tyr(PO(NMe₂)₂)-OH. Further, a phenylalanine and a non-hydrolysable mimic of a phosphorylated tyrosine, 4-phosphonomethyl-L-phenylalanine (pmp), were incorporated during the peptide synthesis. Non-hydrolysable analogues of phosphorylated amino acids can confer long-lasting biological effects that facilitate the studies of possible otherwise labile phosphoproteins. The phosphonate analog of phosphotyrosine (pY) used here, is based upon substitution of the labile bridging oxygen^[23] (Figure 7).

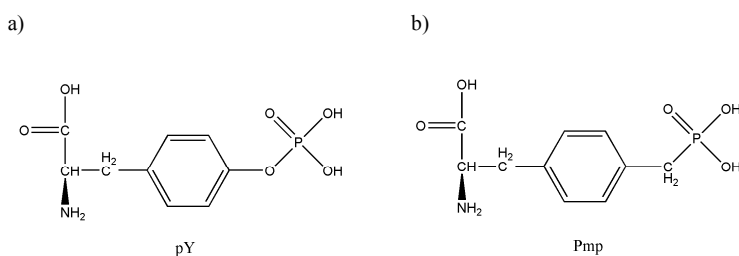


Figure 7: Phosphotyrosine and its non-hydrolysable analogues. (a) Phosphotyrosine (pY), (b) 4-phosphonomethylphenylalanine (Pmp).

The chosen strategy allows the use of the native C₄₈₁ in the ER β LBD for ligation of the synthetic peptide to the protein thioester. The synthesis was started from pre-loaded Fmoc-Gly-Wang resin (Figure 8), followed by cycles of deprotection with piperidine in N,N – Dimethylformamide (DMF) and coupling of the next Fmoc amino acid using N,N'-Diisopropylcarbodiimide (DIC) and 1-Hydroxybenzotriazol (HOBT). The phosphotyrosine residue was hydrolyzed to the corresponding phosphate after peptide cleavage by dissolving the resin-released peptide in a mixture of TFA and H₂O overnight at 4 °C. This protocol allowed the synthesis of the C-terminal part of the ER β LBD with a native tyrosine and with the incorporated modifications pTyr, Phe and Pmp.

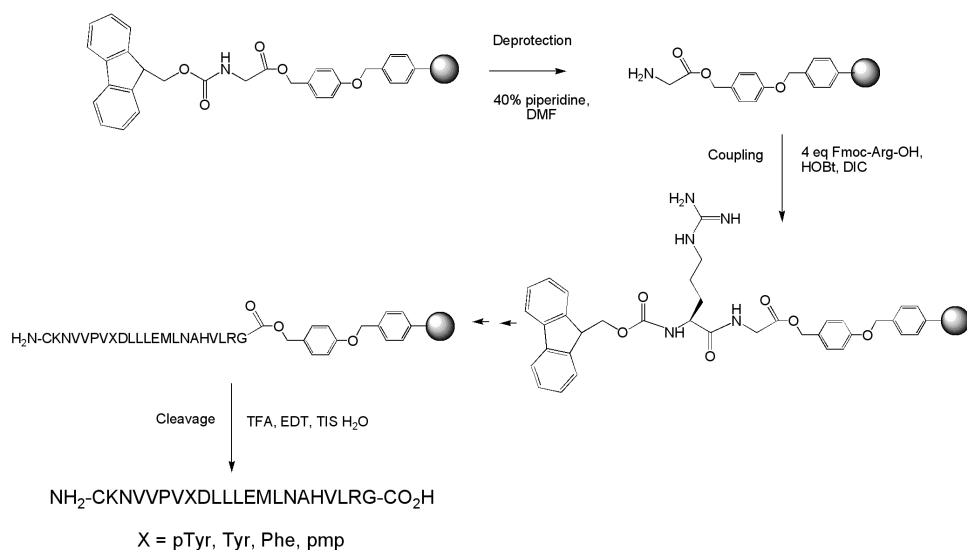


Figure 8: Modular synthesis of the peptides with an N-terminal cysteine using repeated amino acid coupling and Fmoc deprotection steps.

The crude peptides were purified by RP-HPLC on a preparative C18 column. The integrity and purity (>95% at 210 nm) of the peptides was confirmed by analytical RP-HPLC. The estimated yields obtained were 10% for the non-phosphorylated peptide (calcd. mass 2495.34 Da, detected mass 2496.26), 22% for the phosphorylated peptide (calcd. mass 2575.31 Da, detected mass 2576.40 Da), 24% for the Y488F peptide (calcd. mass 2479.34 Da, detected mass 2480.4) and 2% for the pmp peptide (calcd. mass 2557.33 Da, detected mass 2559.2). Due to the extensive preparation and the low yield, the pmp peptide was not further used in the following studies.

The native chemical ligation reaction of the isolated ER β LBD α -thioester and the peptide with an N-terminal cysteine was performed immediately after the purification of both reaction partners. Both building blocks were incubated at 4 °C under non-denaturing conditions *in situ* in the presence of the thiol MESNA (Figure 9). Typically 100% ligation could be achieved by using a 20-fold molecular excess of the peptide. In order to remove non-ligated peptide the samples were extensively dialyzed at 4 °C or more efficiently via size-exclusion chromatography using a FPLC system.

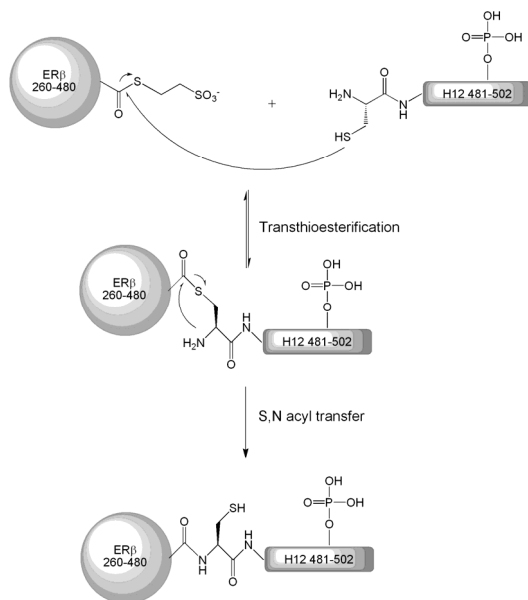


Figure 9: Schematic representation of the native chemical ligation reaction between the truncated ER β LBD α -thioester and the Y₄₈₈ phosphorylated helix 12 peptide with an N-terminal cysteine.

Characterization of the semisynthetic proteins (27.6 kDa) by SDS-PAGE and ESI mass spectrometry (Figure 10) indeed showed the clean conversion of ER β LBD α -thioester (aa 260-480) to the ligated protein including helix 12 (aa 260-502), with no indication of the presence of non-reacted ER β LBD α -thioester. Successful phosphorylation of the LBD was additionally confirmed by immunoblotting using an α -phosphotyrosine FITC conjugate antibody. Only one band corresponding to the phosphorylated ER β LBD could be detected (Figure 10b). Using this approach, both the non-phosphorylated (ER β LBD syn H12 Y₄₈₈) and the phosphorylated (ER β LBD syn H12 pY₄₈₈) ER β LBD could be successfully generated.

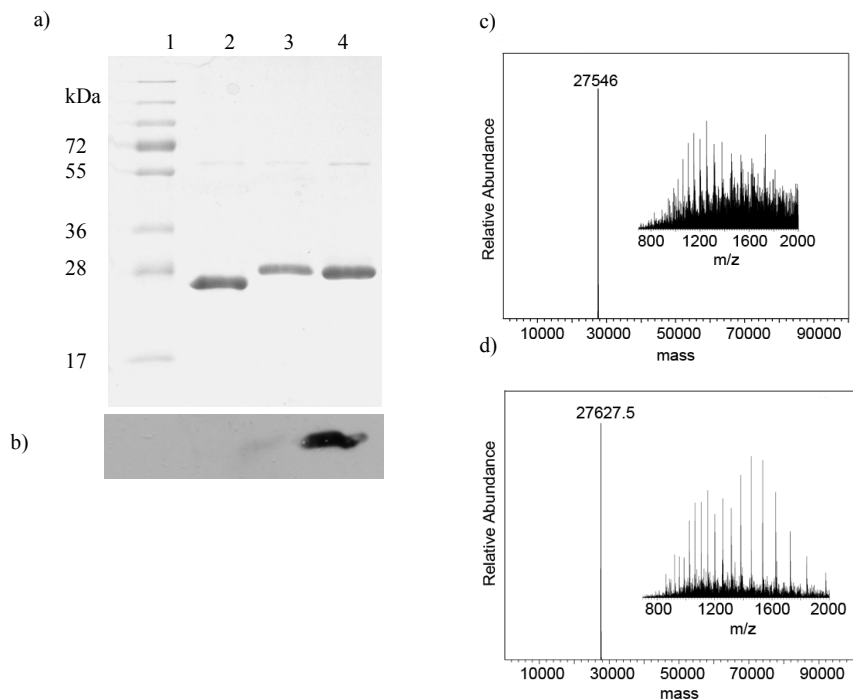


Figure 10: Characterization of the semisynthetic ER β LBDs (aa 260-502); (a) SDS-PAGE gel (15 %; Coomassie stained). Lane 1: molecular weight marker, lane 2: truncated ER β LBD α -thioester, lane 3: non-phosphorylated ER β LBD, lane 4: phosphorylated ER β LBD; (b) Immunoblot using an α -phosphotyrosine antibody; (c) deconvoluted mass spectrum of non-phosphorylated ER β LBD (calcd. mass 27535 Da) and (d) phosphorylated ER β LBD (calcd. mass 27615 Da).

The ligation reactions of the ER β α -thioester and the Y488F peptide and the Pmp peptide, respectively were performed as described for the pY₄₈₈ and Y₄₈₈ peptide. The success of the ligation was confirmed by SDS-PAGE (Figure 11). For comparison, the recombinant non-phosphorylated ER β LBD (aa 260 – 502) was also produced via expression of the complete LBD including helix 12 in *E. coli*. Therefore, the DNA sequence was subcloned in a bacterial vector allowing the heterologous expression of the ER β LBD (aa 260 – 502) with an N-terminal His₆ tag. After native purification via immobilized metal-ion affinity chromatography (IMAC) the His₆-tag could be removed by treatment with the protease thrombin (Figure 11).

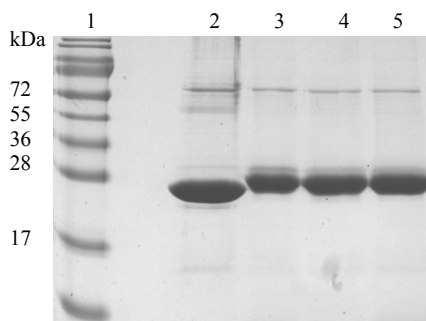


Figure 11: SDS-PAGE gel (15%, Coomassie stained) of ligated modified and heterologous expressed and purified ER β LBD (aa 260 – 502). Lane 1: Molecular weight marker, lane 2: truncated ER β LBD α -thioester, lane 3: Recombinant ER β LBD (aa 260 – 502) with a calculated mass of 27534.7 Da, lane 4: semisynthetic non-hydrolysable phosphorylated ER β LBD using pmp, lane 5: semisynthetic ER β LBD with Y488F.

4.3 Structural Analysis using Circular Dichroism Spectroscopy

The native folding of the semi-synthetic protein constructs was investigated via circular dichroism. Far ultraviolet circular dichroism (CD)^[24] spectral measurements of the semi-synthetic ER β LBD constructs provide a sensitive technique for examining the correct secondary folding and the thermal stability of the proteins^[24]. The CD spectra were measured in phosphate buffer at 10 °C for optimal protein stability. The measurements were performed in the absence and presence of equimolar amounts of the natural agonist estradiol (E₂) and the antagonist *trans*-4-hydroxytamoxifen (HT). The control samples without ligand contained 0.1 % ethanol, as ethanol was used as solvent for E₂ and HT, respectively. The reproducibility of the spectral data was confirmed by repeating each experiment three times.

The CD spectra of all protein constructs featured the typical shape of proteins with a high α -helical content featuring an absorption maximum at 192 nm and two absorption minima at 208 nm and 222 nm (Figure 12) and were comparable to previously reported CD spectra of the ER LBD^[25]. The semi-synthetic approach to generate the ER β LBD thus leads to correctly folded proteins. CD spectra in the presence of both estradiol (E₂) and *trans*-4-hydroxytamoxifen (HT) were similar to the spectrum in the absence of ligand, suggesting that any conformational changes induced by ligand binding do not involve changes clearly observable via CD in the overall secondary structure. This result coincides with previous studies of the effects of ligand binding on the secondary structure of the ER LBD^[25].

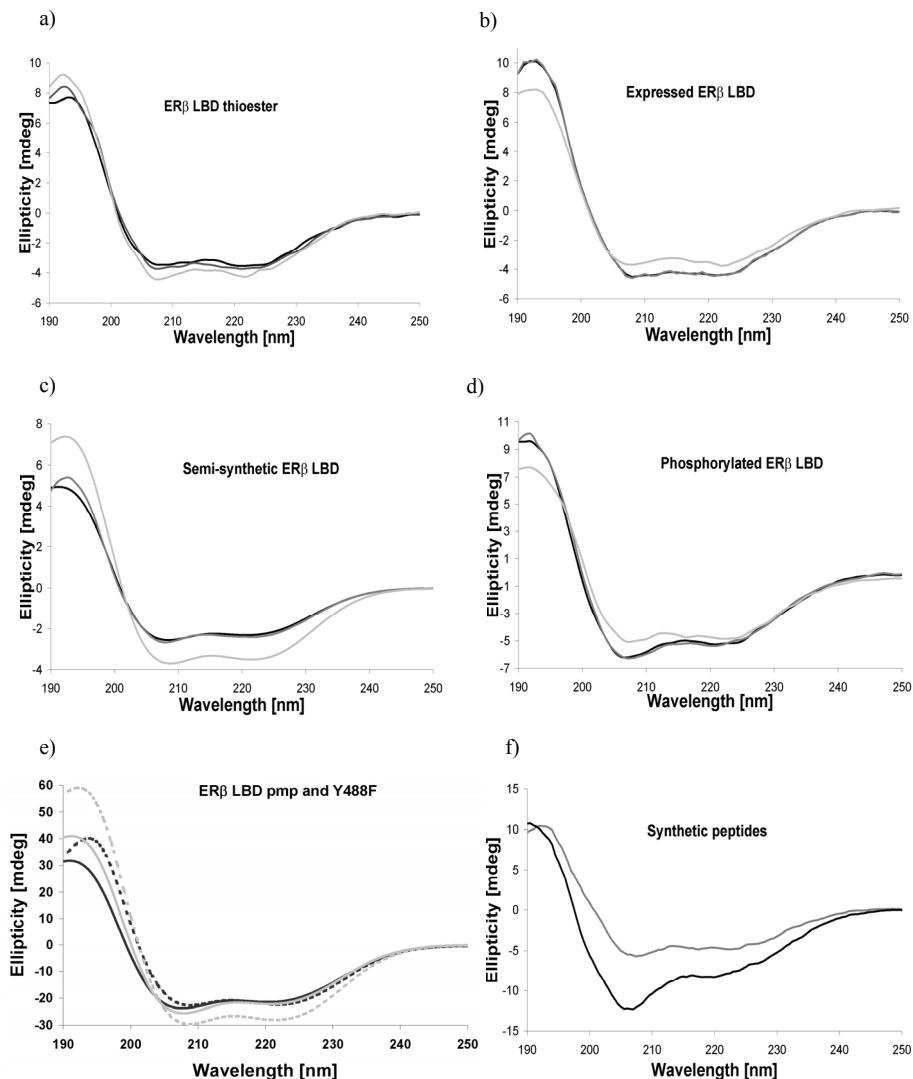


Figure 12: CD spectra of recombinant and semi-synthetic ER β LBDs (a-d) without ligand (black line), with estradiol (dark grey line) and with tamoxifen (light grey line), a) ER β LBD α -thioester, b) recombinant ER β LBD, (c) ER β LBD syn H12, (d) ER β LBD syn H12 pY₄₈₈, (e) ER β LBD syn H12 pmp (continuous line) and ER β LBD syn H12 Y488F (dashed line) and (f) synthetic helix 12 peptide (black line) and synthetic helix 12 phosphopeptide (light grey line).

A quantitative, concentration-independent estimation of the helical character of the peptides can be made by using the ratio $-\theta_{192 \text{ nm}}/\theta_{222 \text{ nm}}$ ^[26]. In Table 1, the values of this ratio are given for all of the ER β LBD constructs. Comparing the data, not only the complete LBDs showed a high α -helical content, but also the truncated ER β LBD α -thioester, lacking helix

12, featured a high α -helical content (an ideal helix features a value of +2.63^[26]) reflecting the stable fold of helices 1-11 in the absence of helix 12.

Table 1: Overview of the helicity of the expressed and semi-synthetic proteins, and the synthetic helix 12 peptides, determined by the $-\theta_{192\text{ nm}}/\theta_{222\text{ nm}}$ ratios.

Protein/ Additive	ER β LBD α - thioester	Expressed ER β LBD	ER β LBD syn H12	ER β LBD syn H12 pY ₄₈₈	ER β LBD syn H12 pmp	ER β LBD syn H12 Y488F	Helix 12 peptide	Helix 12 peptide pY ₄₈₈
no ligand	2.17	2.29	2.13	1.82	1.89	2.29	2.14	1.30
E ₂	2.24	2.29	2.22	1.87	1.96	2.30	-	-
HT	2.16	2.25	2.10	1.58	-	-	-	-

Intriguingly, the phosphorylated constructs (pY₄₈₈ and pmp) featured a somewhat lower α -helical content than the corresponding non-phosphorylated constructs (pY₄₈₈ and F₄₈₈). This was especially true for the helical character of the two synthetic peptides, based on helix 12 (Figure 12 f). Both peptides were measured in the absence of any ligand under the same conditions as for the protein constructs. The non-phosphorylated peptide showed a high $-\theta_{192}/\theta_{222}$ ratio (2.14), typically for a significant α -helical conformation, whereas the phosphorylated peptide featured a lower $-\theta_{192}/\theta_{222}$ ratio (1.30), suggesting a more flexible conformation of helix 12.

Thermal unfolding studies with the different protein constructs were performed to give further insights to the behavior of the semi-synthetic proteins with respect to the stability of the conformation. The thermal stability of the different protein constructs was determined using temperature as the denaturing parameter by measuring the ellipticity at 222 nm as a function of temperature. A representative unfolding profile of the ER β LBDs in the absence and presence of E₂ or E₂ plus a cofactor peptide is presented in Figure 13. The analysis of thermal melting curves was performed by fitting the data to a two-state transition. The midpoint (50%) of unfolding temperature (T_M) values and apparent enthalpies of the unfolding transition are presented in Table 2. The structural stability of the semi-synthetic and expressed proteins is comparable featuring a mean T_M of ~ 50 °C. Addition of E₂ to the ER β LBDs caused in all cases a significant change in the T_M of unfolding (>12 °C). An additional small increase in the T_M could be observed after addition of cofactor peptide to the ER β LBDs with E₂.

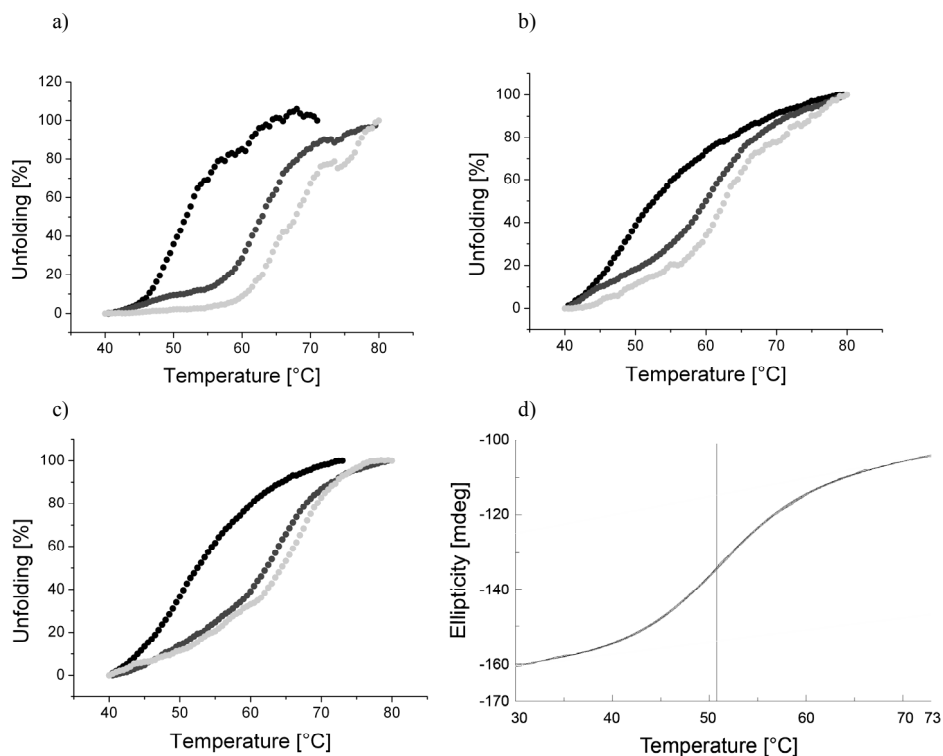


Figure 13: CD melting curves of the expressed and semi-synthetic ER β LBDs in absence of ligand (black), in the presence of E₂ (dark grey) and in the presence of E₂ and cofactor peptide (light grey). The ellipticity at 222 nm was monitored as a function of temperature and correlated to extend of unfolding (a) for the recombinant ER β LBD, (b) for the ER β LBD syn H12 and (c) for the ER β LBD syn H12 pY₄₈₈; (d) example curve (ER β LBD syn H12 pY₄₈₈ without ligand) fitted into a two state model.

Table 2: Effect of E₂ and cofactor peptide (LXXLL) on the thermal unfolding of the expressed and semi-synthetic ER β LBDs

Protein/Additive	Apparent T_M (°C)
Recombinant ER β LBD - ligand	50.13 \pm 0.11
Recombinant ER β LBD + E ₂	62.76 \pm 0.05
Recombinant ER β LBD + E ₂ + LXXLL	64.17 \pm 0.22
ER β LBD syn H12 - ligand	48.34 \pm 0.09
ER β LBD syn H12 + E ₂	61.56 \pm 0.05
ER β LBD syn H12 + E ₂ + LXXLL	61.58 \pm 0.08
ER β LBD syn H12 pY ₄₈₈ - ligand	50.78 \pm 0.06
ER β LBD syn H12 pY ₄₈₈ + E ₂	64.56 \pm 0.07
ER β LBD syn H12 pY ₄₈₈ + E ₂ + LXXLL	67.85 \pm 0.15

Consistent with previous studies, the conformation of the ER LBD can be stabilized by the binding of E₂, most probably by inducing a reorientation of helix 12 to promote the interaction with coactivators^[27]. By binding to the cofactor groove, these coactivator proteins appear to lock the ER LBD in the folded conformation, effectively trapping the ligand in the binding pocket to an even greater degree. Interestingly, the phosphorylated ER β LBD featured a slightly higher stability when bound to E₂ (65 °C), while a significant increase of the unfolding temperature could be observed when E₂ and cofactor peptide are bound to the receptor (68 °C) when compared to the non-phosphorylated reference ER β LBDs. This may indicate that phosphorylation is responsible for the increase in stability of the LBD in complexes with E₂ and the cofactor peptide. Apparently, phosphorylation enhances the binding of the cofactor by further stabilizing the hydrophobic cofactor binding groove. These results raise the possibility that the phosphorylation of the ER β LBD on tyrosine 488 may be involved in cofactor binding. To investigate these effects in greater detail biochemical cofactor recruitment studies were performed.

4.4. Binding Studies using a Cofactor Recruitment On-Chip Assay

NRs recognize so-called LXXLL motifs in coactivator proteins via their AF-2 function on the LBD^[28]. This protein-protein interaction is typically stabilized by an agonistic NR ligand, however it is also operative, with somewhat lower affinity, when the receptor is non-liganded. In order to gain new insight into the role of the Y₄₈₈ phosphorylation with respect to these protein-protein interactions, cofactor binding studies were performed using libraries of peptides representing the important binding epitopes of the cofactor proteins. This microarray technique requires a fluorophore for visualizing the binding event. Similar to the strategy described above (Chapter 4.2) the phosphorylated ER β LBD, as well as the semi-synthetic and recombinant non-phosphorylated reference proteins were generated with an N-terminal His₆-tag. The excess of peptide during the ligation reaction could be removed by immobilized metal-ion affinity chromatography (IMAC). The recombinant reference ER β LBD could directly be purified using IMAC. The integrity and purity of the three proteins was proven by SDS-PAGE (Figure 14).

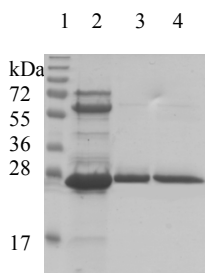


Figure 14: SDS-PAGE gel (15%, Coomassie stained) of the purified ER β LBD with an N-terminal His-tag. Lane 1: Molecular weight marker, lane 2: Recombinant His₆-ER β LBD (calcd. mass: 29698 Da), lane 3: His₈-ER β LBD syn H12 (calcd. mass: 29281.8 Da), lane 4: His₈-ER β LBD syn H12 pY₄₈₈ (calcd. mass 29361.8 Da).

A first set of binding experiments was performed using a library of 53 cofactor peptides of ~25 amino acid residues, representing both coactivator (LXXLL) and corepressor (LXXXIXXXL) proteins. These peptides were immobilized on a porous metaloxide carrier. In individual experiments, both the phosphorylated and non-phosphorylated proteins were pumped several times through the three-dimensional metaloxide carrier chip^[29] at 25 °C (Figure 15). Subsequently, the binding of the proteins to the specific peptides was assessed by hybridizing the protein constructs with a fluorescently labeled antibody against the His-tag on the LBD. Each experiment was performed in three separate experiments to verify the measured data. During each experiment, images were collected at different incubation cycles, clearly indicating a time-dependent increase of binding. As control, a reference membrane was incubated without protein to subtract unspecific binding of the antibody. The obtained peptide binding profiles are reported in Figure 16.

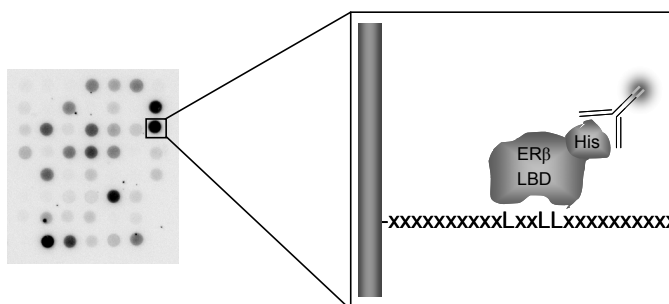


Figure 15: Estrogen receptor-coregulator-interaction profiling. A selected set of 53 coregulator peptides immobilized on a porous metaloxide carrier were incubated with non-phosphorylated and phosphorylated ER β LBDs and binding was detected by fluorescent α -His antibody. Each spot represents distinct immobilized coregulators. The intensity of each black spot resembles fluorescence intensity and thus the amount of bound protein.

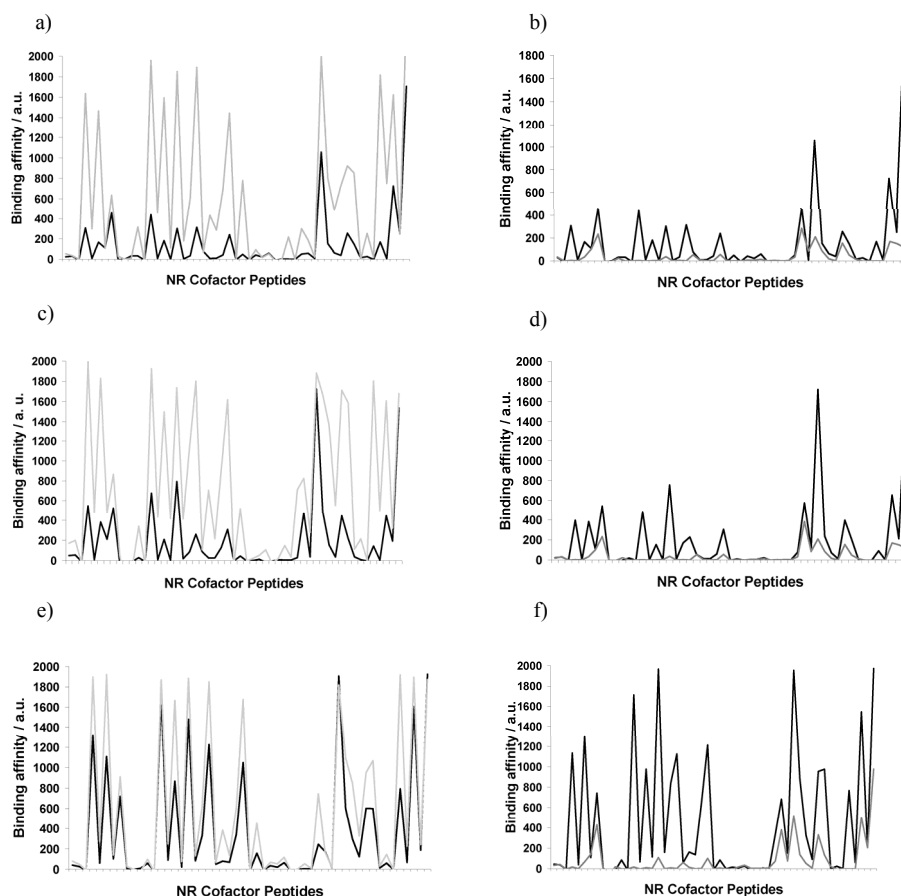


Figure 16: Peptide binding profiles of the recombinant ER β LBD (a and b), the ER β LBD syn H12 (c and d) and ER β LBD syn H12 pY₄₈₈ (e and f) to a set of 53 cofactor peptide sequences. The binding was analyzed in presence or absence of E₂ (left) and HT (right). Protein without ligand is shown in dark grey and protein with ligand (E₂ or HT) is shown in light grey. The reference proteins feature similar binding profiles, while the phosphorylated ER β LBD showed an increased affinity to the peptides in the absence of ligand.

All three ER β constructs showed relatively similar peptide binding profiles in the presence of E₂ and HT, respectively. Typically, peptides featuring high affinity to the non-phosphorylated construct also bind strongly to the phosphorylated protein. Nevertheless, for selected peptides a slight differentiated binding occurs, as can be deduced from the relative affinities amongst the peptides. A significant difference between the two reference proteins and the phosphorylated LBD was observed in the non-liganded state. The non-phosphorylated constructs both featured a significantly lower binding affinity to all peptides when estradiol was not present. The phosphorylated construct, *vice versa*, showed only minimally decrease

in binding in the absence of an agonist for most of the peptides tested. The ligand independent binding of the ER β LBD is thus increased upon phosphorylation. Treatment of the proteins with the antagonist *trans*-4-hydroxytamoxifen resulted in a strong overall lowering of the peptide affinity in all three proteins.

It could thus be demonstrated that the semi-synthesis of the ER β LBD did not impair the cofactor binding activity when compared to the native, expressed, LBD. Furthermore, phosphorylation of tyrosine 488 resulted in an enhanced binding affinity to all measured cofactor peptides in the absence of a ligand, although it had no effect on cofactor binding when the receptor was bound to E₂ and HT. It should be noted that in this study the peptides are immobilized on a surface which leads to high local concentrations of peptide and possibly to concomitant multivalency effects, similar to the natural cofactors.

4.5 Binding Studies using a Cofactor Recruitment FRET Assay

Cofactor recruitment studies in solution enable the investigation of individual binding events between the ER β LBD and coactivator peptides and prevent the possibility of multivalency effects due to high local concentrations of cofactor peptides as observed for peptides immobilized on surfaces. The binding of a selected set of coactivator peptides to the non-phosphorylated (semi-synthetic and recombinant) and phosphorylated (semi-synthetic) constructs was analyzed using a TR-FRET (time-resolved fluorescence resonance energy transfer) based binding assay^[30] in solution. In this assay, a fluorescent labeled streptavidin binds to a biotinylated cofactor peptide that, in turn, must be recruited by the His-tagged ER β LBD to bring the fluorophores allophycocyanin (APC) and europium (Eu) within close proximity (Figure 17). Subsequent excitation of Eu with light at a wavelength of 340 nm leads to an energy transfer that could be determined by measuring the emitted light of APC at 665 nm.

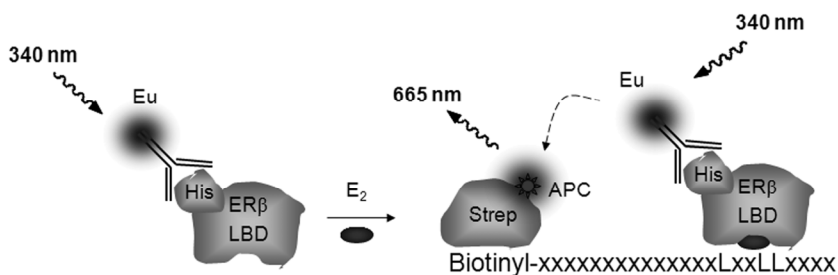


Figure 17: Illustration of the *in vitro* cofactor recruitment assay. Binding of an agonist to the His-tagged ER β LBD will recruit a biotinylated coactivator peptide. Eu-labeled α -His antibody and streptavidin-conjugated APC will assemble into the complex resulting in FRET.

Using this peptide recruitment assay, we analyzed 31 different coactivator peptides both in absence and presence of the ligands E₂ and HT. Increasing concentrations of cofactor peptides were incubated with a mixture of Eu-labeled α -His antibody, ER β LBD, streptavidin-conjugated APC and if necessary, with ligand over night at 4 °C, and FRET signals were subsequently measured. Plotting of the response data against the peptide concentration yielded sigmoidal curves, which allowed the determination of EC₅₀ values of peptide binding. As an example, the binding of the different protein constructs to four selected peptides, which exhibit significant binding to all three proteins in the presence of estradiol, is displayed (Figure 18). Peptides that exhibit high affinity for the recombinant reference protein also showed a strong binding to the semi-synthetic ER β LBDs. E₂ increased binding of the peptide in all cases and HT, conversely, decreased or rather suppressed binding of the cofactor peptides. The high binding affinities, observed here, confirmed the correct folding and cofactor binding activity of the semi-synthetic proteins.

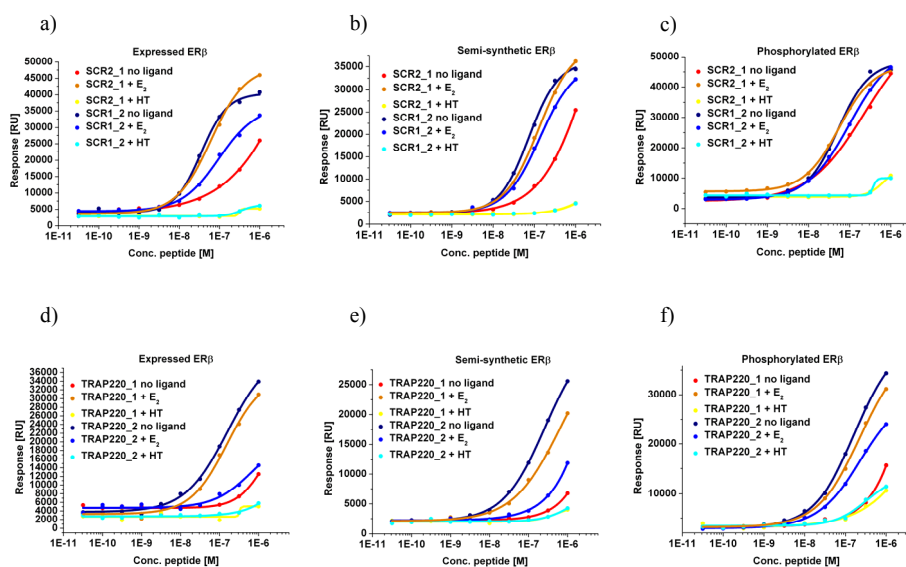


Figure 18: Exemplary ER β LBD cofactor peptide recruitment profiles influenced by different ligands. Each graph shows the binding profile of two distinct peptides (blue and reddish colors) in the absence of ligand and in the presence of estradiol (E₂) and *trans*-4-hydroxytamoxifen (HT), respectively; (a and d) peptide binding to the recombinant ER β LBD; (b and e) peptide binding to the ER β LBD syn H12; (c and f) peptide binding to the ER β LBD syn H12 pY₄₈₈. All protein constructs featured similar cofactor binding profiles.

In order to monitor a potential difference in cofactor recruitment due to the phosphorylation, the binding pattern of two exemplary cofactor peptides for the reference ER β LBDs and the phosphorylated ER β LBD were compared in the absence and presence of ligands (Figure 19). The resulting EC₅₀ values, as far as determinable, of the cofactor binding to the three ER β LBDs were summarized in Table 4.

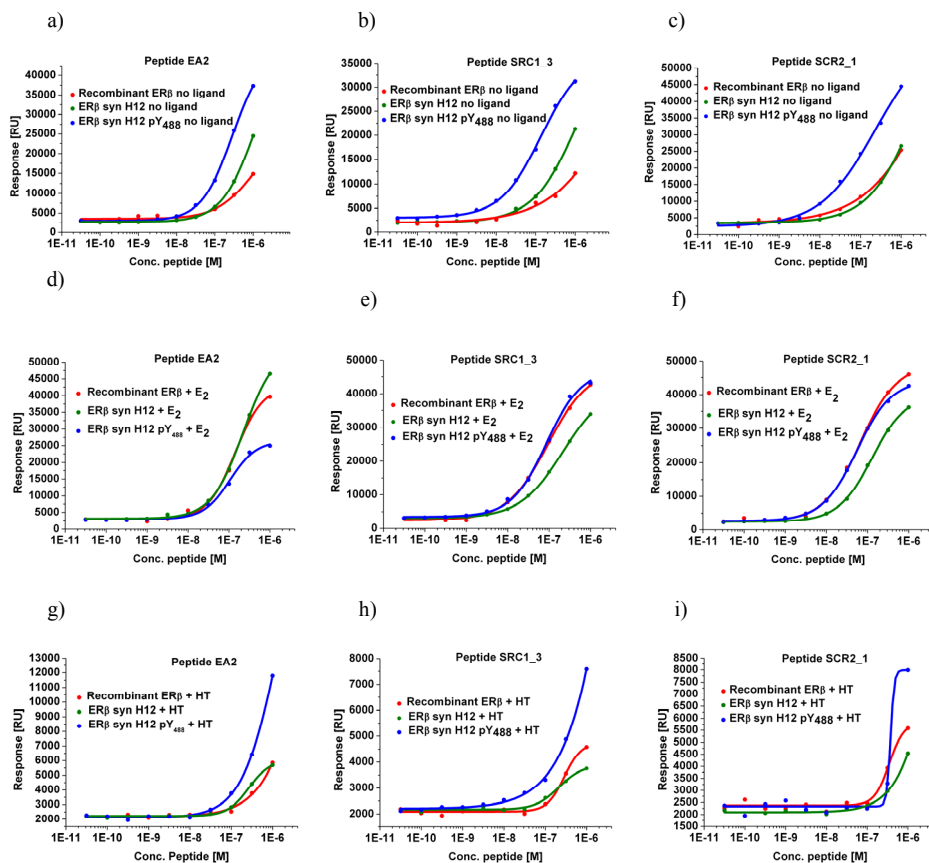


Figure 19: Exemplary binding patterns of a selected set of cofactor peptides arranged per protein in the absence of ligand (a - c) and in the presence of estradiol (d - f) and *trans*-4-OH-tamoxifen (g - i). With estradiol, all three recombinant ER β LBD (red), ER β LBD syn H12 (green) and ER β LBD syn H12 pY₄₈₈ (blue) featured similar binding profiles, whilst the efficiency of the FRET signal without ligand was highly increased for the phosphorylated protein. HT also led to a slight increase in the FRET signal for the phosphorylated ER β LBD.

While the two reference proteins feature comparable cofactor binding patterns in the presence and in the absence of ligand, the phosphorylated ER β LBD shows several differences in cofactor binding. The most significant difference can be observed in the absence of ligand (Figure 19a-c). For almost all measured cofactor peptides, tyrosine

phosphorylation results in a strong increase in the FRET signal. This increase in FRET efficiency is accompanied with a remarkable (up to 10-fold) increase in cofactor binding affinity (italics in Table 4) and may reflect conformational changes in the ER β – cofactor interaction due to the phosphorylation. These results are in line with the observed increase in cofactor binding for the phosphorylated ER β LBD in the on-chip studies. A similar effect could be detected for the phosphorylated ER β in the presence of the antagonist hydroxytamoxifen. The reference proteins show the expected decrease in the FRET signal and the EC₅₀ values (Figure 19g-i; Table 4). Addition of HT to the phosphorylated protein as well features a decrease in the FRET signal. However this decrease was not as remarkable as for the non-phosphorylated ER β LBD. The results observed for the estradiol-bound ER β LBD highly differ from the results of the apo-protein and the antagonist occupied ER β LBD. The cofactor binding pattern, if at all, only slightly varied from the binding pattern of the references (Figure 19d-f). In most cases there is no effect of phosphorylation detectable (Figure 19e-f), but in a few cases, the cofactor binding to the phosphorylated ER β LBD is slightly increased or decreased (Figure 19d). The relative mildness of the differences observed in these biochemical studies is not surprising as, for example, the binding of E₂ typically only results in an around a 10-fold increase in peptide binding versus the apo-form. It is well established that small changes observed in biochemical studies can result in large effects on the cellular level, since these differentiating effects are subsequently amplified via the gene transcription mechanism.

Table 4: EC₅₀ values in mol/L [M], of the different cofactor peptides for the two control ERβ LBDs and for the phosphorylated ERβ LBD in the absence of ligand and in the presence of E₂. An increase of FRET efficiency for the phosphorylated protein is marked in *italics* and a decrease in the signal is highlighted in **bold**.

Peptide	Expressed ERβ LBD		Semi-synthetic ERβ LBD		Phosphorylated ERβ LBD	
	no ligand	E ₂	no ligand	E ₂	no ligand	E ₂
SCR2_1	--	6.15E-08	--	1.26E-07	<i>2.26E-07</i>	<i>5.5E-08</i>
SCR1_2	9.99E-08	3.48E-08	1.35E-07	7.3E-08	<i>9.98E-08</i>	5.2E-08
SCR2_3	--	3.17E-08	7.55E-08	5.89E-08	2.14E-07	7.64E-08
SHP_1	3.02E-07	2.82E-07	--	3.03E-07	--	--
SHP_2	--	2.83E-07	--	--	--	--
D22	6.1E-08	3.79E-08	1.14E-07	9.89E-08	1.88E-07	9.57E-08
D47	--	6.68E-08	--	2.1E-07	--	--
DAX3	2.08E-07	5.12E-08	1.09E-07	8.58E-08	2.64E-07	1E-07
RIP140_5	1.31E-07	5.11E-08	1.78E-07	1E-07	1.95E-07	8.06E-08
RIP140_6	--	5.23E-08	--	1.15E-07	--	1.41E-07
RIP140_8	1.68E-07	7.94E-08	2.7E-07	2.05E-07	--	1.53E-07
RIP140_9	--	6.17E-08	3.05E-07	1.17E-07	--	1.72E-07
CBP_1	--	2.33E-07	--	--	--	--
PGC_1	--	4E-08	--	1.08E-07	--	2.55E-07
RIP140_3	--	--	--	--	--	--
HRCoA2	--	--	--	--	--	--
TRAP220_1	--	1.52E-07	--	--	--	2.31E-07
TRAP220_2	--	1.58E-07	--	2.6E-07	<i>2.42E-07</i>	1.67E-07
SRC3_1	1.89E-07	8.17E-08	--	1.3E-07	<i>5.98E-08</i>	9.43E-08
SRC3_2	--	4.09E-08	2.73E-07	8.21E-08	<i>7.97E-08</i>	6.61E-08
SRC1a_4	1.28E-07	3.89E-08	1.62E-07	7.76E-08	<i>8.81E-08</i>	5.19E-08
SRC1_3	--	9.35E-08	--	2.63E-07	<i>1.39E-07</i>	<i>8.64E-08</i>
RIP140_1	--	1.39E-07	--	1.18E-07	<i>2.45E-07</i>	<i>9.91E-08</i>
RIP140_7	--	1.61E-07	--	1.91E-07	--	--
D30	1.28E-07	5.19E-08	2.18E-07	1.31E-07	1.17E-07	1.09E-07
TIP60	--	--	--	--	--	--
PERC-1	--	4.17E-08	--	1.17E-07	--	1.39E-07
PERC-2	--	2.41E-07	--	--	--	--
EA2	--	1.43E-07	--	2.19E-07	<i>2.93E-07</i>	<i>1.05E-07</i>
ARAF1_4	--	--	--	--	--	--
SRC1_1	--	9.72E-08	--	2.38E-07	--	--
buffer	--	--	--	--	--	--

These results are consistent with a model in which the ERβ wild type in the absence of ligand adopts an inactive ligand pocket conformation. When estradiol is bound, the wildtype ERβ adopts an active conformation, induced by a conformational shift of helix 12 against the surface of the LBD that traps the ligand inside the pocket. This conformation enables the generation of a hydrophobic groove on the surface of the LBD that can be recruited by coactivators. By contrast, phosphorylation on tyrosine 488 appears to shift the equilibrium of conformations, attainable by the LBD, to a more agonist like state, whether ligand is bound or not and thus enhances the affinity for cofactor peptides in the absence of E₂ (Figure 20). Further, phosphorylation seems to partially abolish the antagonistic effect of HT by slightly

enhancing the cofactor recruitment. Phosphorylation might permit helix 12 to take a position that enables the formation of the cofactor binding groove without interfering with the bulky antagonist. Most probably, phosphorylation of tyrosine 488 mimics the agonist-bound pocket conformation by stabilizing interactions with nearby regions of the ER LBD in the apo-protein or in the presence of HT.

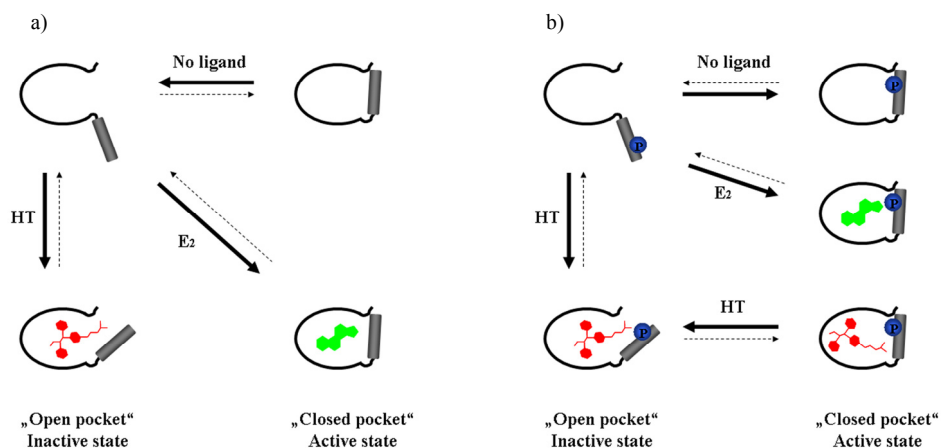


Figure 20: Schematic model for the ER β , illustrating the potential process of conformational reorganization, involved in ligand association and dissociation in wild type ER β (a) and ER β phosphorylated at Y₄₈₈ (b). The open pocket conformation is unoccupied and inactive with respect to coactivator binding. The closed pocket conformations are active and enable the ER β to recruit coactivators. The principal relative rates of isomerisation between the various states in absence of ligand and in presence of estradiol (E₂) and 4-hydroxytamoxifen (HT) are illustrated by the thickness of the arrows connecting the states. The wildtype ER β favors the inactive open pocket conformation in absence of ligand and the active conformation in presence of E₂. HT keeps the ER β in an inactive state. The phosphorylated ER β favors the active closed pocket conformation in absence of ligand and in the presence of E₂ and is able to create the active conformation in the presence of HT.

4.6 Co-Crystallization Studies of the Phosphorylated ER β LBD in Complex with Estradiol and Cofactor Peptide

To confirm the exact conformational structure of the phosphorylated ER β LBD and to get deeper insights into how this potential conformational change influences the function of the receptor with respect to interactions with other proteins, the defined three-dimensional crystal structure of the semi-synthetic post-translationally modified LBD in complex with E₂ and a cofactor LXXLL peptide was investigated by X-ray co-crystallography.

One of the most important requirements for achieving appropriate protein crystals is the preparation of a suitable protein construct with a high degree of purity. Various x-ray

structures of non-phosphorylated NR LBDs, including ER β LBDs, have been successfully dissolved in the presence of different agonists, antagonists and coactivator peptides in the past^[31]. In order to get the ER β LBD with a high purity, a strategy reported by Alvarez *et al.* was used^[31b]. The sequence used for the ER β LBD, ranging from L₂₆₀ to G₅₀₂, was exchanged by the following sequence from MD[D₂₆₁ – L₅₀₀]DD. The DNA for the completely expressed LBD was subcloned into a vector expressing the ER β LBD featuring the above sequence without any tag. Analogously the ER β LBD α -thioester (aa MDD₂₆₁-K₄₈₀) was generated as described earlier via intein-catalyzed thiolysis without ligand. After peptide synthesis and purification of the sequence NH₂-CKNVVPVpYDLLLEMLNAHVLDLDD-CO₂H with a phosphorylated tyrosine (calcd. mass: 2593.93 Da, detected mass: 2593.6 Da; purity >95%), both building blocks underwent chemical ligation and the resulting protein was purified via estradiol affinity chromatography. After carboxymethylation of the bound protein, the elution of the purified ER β LBD was performed with increasing amounts of estradiol. Size exclusion chromatography (SEC) was carried out and the excess of ligand was finally be removed by passing the solution through a G-25 desalting column. For the expressed reference ER β LBD, the clarified *E. coli* lysate was directly applied to the estradiol affinity column without prior purification steps. The purity of both proteins was confirmed by SDS-PAGE and SEC (Figure 21).

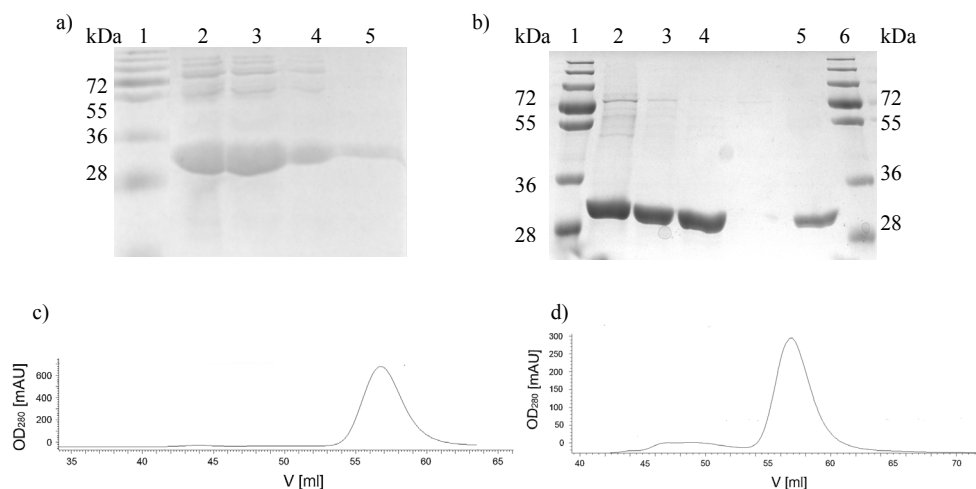


Figure 21: Analysis of the purification of the expressed and the semi-synthetic phosphorylated ER β LBD; (a) SDS PAGE gel (15%, Coomassie stained) of expressed and (b) phosphorylated ER β LBD. Lanes 1 and 6: Molecular weight marker, lane 2: ER β LBD after elution from the estradiol affinity column using 50 μ M E₂, lane 3: using 100 μ M E₂ and lane 4: using 200 μ M E₂ and lane 5: ER β LBD after size exclusion chromatography; (c) Elution profile of expressed ER β LBD (calcd. mass: 27553.5 Da) and (d) phosphorylated ER β LBD (calcd. mass: 27633.5 Da) after SEC-analysis using FPLC.

Both proteins were concentrated to approximately 12 mg/ml and combined with the cofactor peptide containing an LXXLL motif (NH₂-LTERHKILHRLLQEGSPSD-CO₂H) at a molar ratio of 1.5:1 peptide : protein-ligand complex. The synthesis of the peptide was performed via SPPS using Fmoc protected Rink amide MBHA resins. The integrity of the peptide was confirmed by HPLC (calcd. mass: 2229.52 Da, detected mass: 2228.54) and after purification the product was obtained in a yield of 28% and a purity of more than 99%. The coactivator peptide was extensively dialyzed for two days at 4 °C against the protein buffer. This step is essential to remove the trifluoroacetic acid (TFA) salt used for the purification of the peptide as it could negatively influence the crystallization behavior of the protein. The crystallization was subsequently performed for the non-phosphorylated ER β LBD with the sitting drop vapor diffusion method using initial commercial screens with a drop volume of 200 nl at 20 °C and 4 °C. These so called *Sparse Matrix Screenings* ^[32] give information about the crystallization behavior of the protein under specific conditions.

Initial crystals grew in several conditions generally containing 20% of the precipitant polyethylenglycol (PEG). Reproduction of these crystals for fine screening experiments were performed with the hanging drop vapor diffusion method using drop volumes between 2 and 5 μ l. Optimal crystals could be observed for the non-phosphorylated ER β LBD/E₂/cofactor peptide complex in a drop containing a 1: 0.5 ratio mixture of protein-ligand-peptide solution to reservoir solution of 0.2 M Ammonium Acetate, 0.1 M BIS-TRIS pH 5.5 and 17% PEG10000 (v/v) at 4 °C. The protein crystals were grown for five days before reaching their final size of 0.5 x 0.4 x 0.3 mm (Figure 22). Prior to data collection, all crystals were briefly soaked in a solution containing the mother liquor and 20% glycerol (v/v) before flash-frozen in liquid nitrogen. X-ray diffraction data of a recombinant non-phosphorylated ER β LBD/E₂/peptide crystal were collected at a temperature of 100 K using a MAR dtb detector at the Nonius AXS MICRO Star. The maximal resolution for the crystal of the non-phosphorylated ER β LBD complex adds up to 2.2 Å.

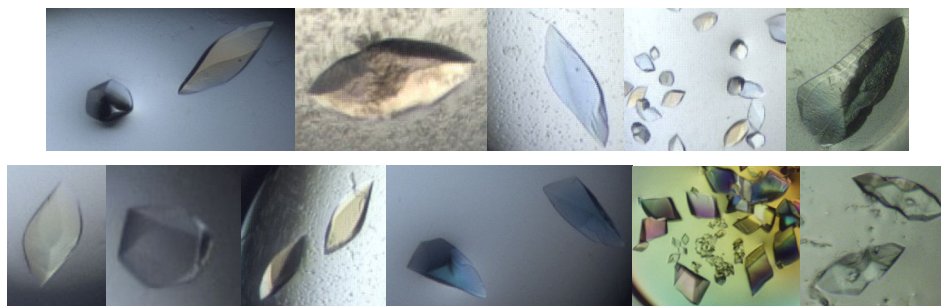


Figure 22: Co-Crystals of the expressed ER β LBD (upper row) and phosphorylated ER β LBD (lower row) in complex with E₂ and cofactor peptide under conditions of 0.2 M Ammonium Acetate, 0.1 M BIS-TRIS pH 5.5 and 17% PEG10000 (v/v) at 4 °C.

Due to the successful crystallization of the non-phosphorylated reference protein, co-crystallization of the phosphorylated ER β LBD was performed using the same conditions. High quality crystals could be observed in the same solution as for the reference protein and diffraction resulted in the maximal resolution of 2.2 Å. To get an even higher resolution, the data set of a phosphorylated ER β LBD/E₂/cofactor peptide complex crystal was recorded at 100 K at *the Swiss Light Source* (PSI, beamline PX10SA, Villingen, Switzerland) using a MAR225 CCD detector to a maximal resolution of 1.5 Å.

All data sets were processed and indexed using XDS^[33] and scaled using XSCALE^[33]. In both cases hexagonal crystals could be observed that belong to the monoclinic space group P3₁. The unit cell of the reference crystal had an edge length of $a = 71.9$ Å, $b = 71.9$ Å and $c = 113.3$ Å with the cell angles $\alpha = \beta = 90^\circ$, $\gamma = 120^\circ$. The crystals of the phosphorylated ER β LBD complex featured the same unit cell dimensions. Further crystallographic data are summarized in Table 5.

Table 5: X-ray data collection statistics for ER β LBD complex crystals (non-phosphorylated and phosphorylated)

protein	reference ER β LBD complex	phosphorylated ER β LBD complex
maximal resolution (Å)	2.2	1.5
space group	$P3_1$	$P3_1$
unit cell dimensions	$a = 71.9 \text{ \AA}$ $b = 71.9 \text{ \AA}$ $c = 113.3 \text{ \AA}$ $\alpha = \beta = 90^\circ, \gamma = 120^\circ$	$a = 71.9 \text{ \AA}$ $b = 71.9 \text{ \AA}$ $c = 113.3 \text{ \AA}$ $\alpha = \beta = 90^\circ, \gamma = 120^\circ$
number of molecules per asymmetric unit	2	2
wavelength (Å)	1.548	1.000
completeness (%)	98.6	99.8
R_{sym} (%)	27.9	37.7
mean $1/\sigma$ (I)	14.18	17.55

$R_{\text{sym}} = \sum | I(h)j - \langle I(h) \rangle | / \sum I(h)j$, where $I(h)j$ is the scaled observed intensity of the i^{th} symmetry related observation of the reflexes h and $\langle I(h) \rangle$ is the mean value.

The crystal structures were solved by molecular replacement using the program PHASER^[34]. The phosphorylated ER β LBD complex structure was solved using the ER β LBD complex structure of the recombinantly expressed reference protein as search model. This reference structure was in turn solved in-house by molecular replacement, using the ER β LBD complexed with an agonist and coactivator peptide as a model (pbd code: 1u9e)^[31b]. Cysteine modifications and some flexible loop residues were not included in the model due to poor electron density. The final models of both ER β LBD complex crystal structures were in good agreement with those previously reported^[31a], containing a homo dimer, with the ligand E₂ and coactivator peptide bound to each monomer, and a variable number of water molecules. Each LBD displayed the canonical NR fold composed of 12 α -helices. Five residues prior to serine 264 at the N-terminus of helix 1 and five residues after leucine 495 at the C-terminus of helix 12 are invisible in the electron density and have not been modeled. The coactivator binding site of each LBD monomer is occupied by the coactivator peptide. The binding mode of the used SRC 1 Box 2 peptide is identical to that observed in analogous complexes between related coactivator LXXLL motifs and NR LBDs^[35]. Although the electron density for the coactivator peptide is continuous, not the entire peptide motif was visible, but only the ten residues that correspond closely to the minimal core sequence capable of binding to the cofactor binding groove (RHKILHRLQ).

An overlay of both crystal structures featured the identical arrangement of molecules within the ER β LBD complex (Figure 23). However, for the semi-synthetic ER β LBD, the electron density clearly indicated the phosphorylation of tyrosine 488. This tyrosine, located in the loop between helix 11 and helix 12, exists at the surface of the protein so that the attached phosphate is protruding out from the ER β LBD surface. This orientation of the phosphorylated tyrosine is ideal to enable it to be easily accessed by other proteins, thus allowing for specific protein-protein interactions with the ER β LBD. In previously reported cellular studies phosphorylation on tyrosine 488 had been identified to be responsible for the interaction of the SH2 domain of c-Src with the ER β LBD, resulting in the activation of the c-Src kinase function^[10].

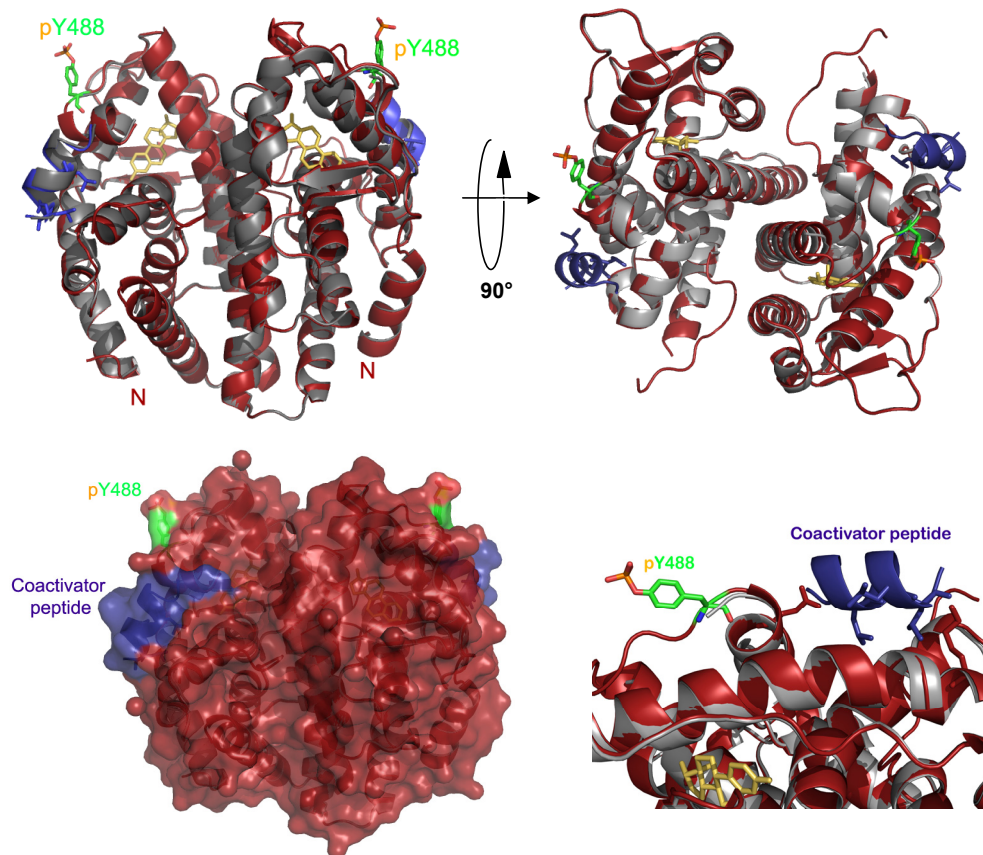


Figure 23: Overall view of the crystal structure of the phosphorylated ER β LBD (red) complexed with cofactor peptide (blue) and estradiol (yellow) and overlaid with the non-phosphorylated ER β LBD (grey). Shown is the ER β LBD homo dimer in two different perspectives resulting from a rotation of 90° over the x-axis. The lower figures show the solid phase representation of the phosphorylated ER β LBD complex (left) and a detail of the phosphorylation region (right). The phosphorylated residue (orange) sticks out of the surface.

In contrast, the solved crystal structure of the agonist-bound ER β LBD gave no further information regarding the possible participation of the phosphorylation on the interaction with coactivator proteins. Together with the results obtained of the cofactor recruitment studies, the crystal structure indicates that the phosphorylated ER β LBD behaves like the wild type ER β in the presence of estradiol. The observed effect of tyrosine phosphorylation on the cofactor binding in absence of ligand or in the presence of hydroxytamoxifen could be accompanied with a conformational change in the LBD, and especially in helix 12. This assumption raises the necessity to determine the crystal structure of an ER β LBD – cofactor peptide complex in the absence of ligand or in the presence of HT. To date, the only solved crystal structure of a nuclear receptor without a ligand is limited to the RXR and some orphan receptors for which no ligand could be identified so far. The difficulty of crystallizing a NR as an apo-protein is probably due to a lack of stability of helix 12. CD thermal unfolding studies as well as biochemical cofactor recruitment assays indicate an enhanced stability of the ER β LBD phosphorylated on tyrosine 488, which promises a higher chance to crystallize the phosphorylated ER β LBD in the absence of ligand. However, upon intensive screening of many different commercial crystallization conditions at 4 °C and at 20 °C, crystals could neither be observed in the presence of HT nor in absence of ligand. Alternative possibilities to determine a structural change in the phosphorylated ER β LBD without a ligand might include NMR studies using N¹⁵ labeled amino acids in the synthetically-produced helix 12. Due to variations in the resonance signals of the labeled amino acids, it would be possible to determine potential distinct dynamics in the phosphorylated and the non-phosphorylated ER β LBD.

4.7 Binding Studies using Surface Plasmon Resonance (SPR)

The role of ER β tyrosine phosphorylation in the activation of the kinase c-Src via recruitment of its SH2 domain^[10] is the current objective of intensive research using cellular assays^[10]. As the exact binding mechanism of c-Src to the ER β LBD via the phosphorylated tyrosine 488 is still unknown on the molecular level, biochemical recruitment techniques using the isolated SH2 domain might give further insights into how phosphorylation enables SH2 binding. In order to investigate a potential binding of the SH2 domain of c-Src to the ER β LBD via the phosphorylated tyrosine 488 *in vitro*, surface plasmon resonance (SPR)^[36] was performed. Studies were carried out using biosensor chips consisting of a gold substrate coated with a non-fouling dextran layer bearing carboxymethyl groups. For this purpose, the sequence of the SH2 domain (aa 151 – 248) and the SH2/SH3 domain (aa 84 – 248) of c-Src were each

cloned into the pTriEx-2-based vector pOPINF^[37] with an N-terminal cleavable His₆-tag. After expression in *E. coli* and purification via IMAC, the His₆-tag was cleaved via a precision protease followed by removal of the cleaved tag and the protease using size exclusion chromatography. The purity of the resulting proteins was confirmed by SDS-PAGE (Figure 24).

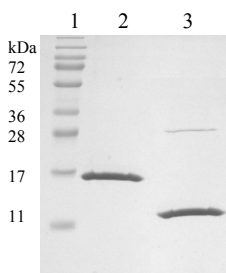


Figure 24: SDS-PAGE gel (15%; Coomassie stained) of the purified SH2 domain of c-SRC. Lane 1: Molecular weight marker; Lane 2: Purified SH3/SH3 domain (calcd. mass 18711 Da); Lane 3: Purified SH2 domain (calcd. mass 11284 Da).

The protein featuring the SH2 domain alone turned out to be quite unstable as it started to aggregate after defrosting or storage at 4 °C. It was therefore decided to select the protein featuring the SH2 and the SH3 domain of c-SRC for immobilization on a biosensor surface via EDC/NHS activation. After regeneration of the surface to remove non-covalently bound material, a solution of the phosphorylated ER β LBD was flown over the sample surface and over a reference surface where no protein was immobilized with increasing concentrations. The subtracted signals are shown in Figure 18. Increasing concentrations of the phosphorylated protein generated an increase in signal. However, incubating the surface with 10 μ M of the non-phosphorylated ER β LBD reference caused an even higher response signal in comparison to the same concentration of phosphorylated protein (Figure 25). This indicates that the observed signal increase is not due to specific binding of the phosphorylated tyrosine 488 to the SH2 domain.

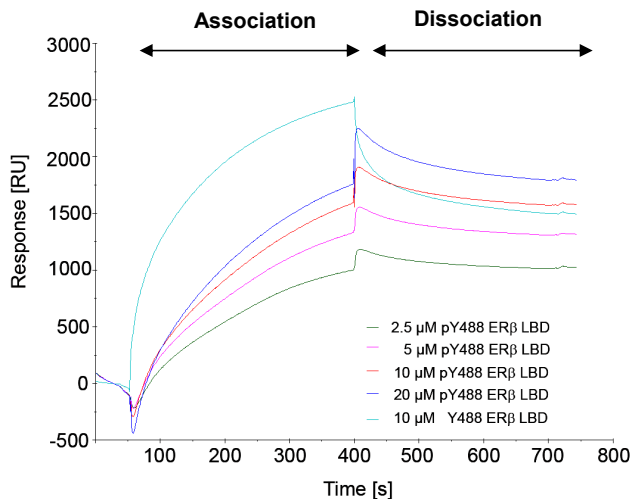


Figure 25: Sensogram of the interaction of the immobilized SH2/SH3 protein with different concentrations of the phosphorylated ER β LBD and with 10 μ M of the non-phosphorylated ER β LBD using SPR. The non-phosphorylated LBD binds even stronger to the SH2 domain than the phosphorylated protein.

One disadvantage of the used immobilization strategy is the heterologous surface due to several lysines in the protein that lead to different orientations of the SH2 domain. Therefore not each binding site of molecule is available for the phosphorylated tyrosine. Unspecific binding of the ER β LBD to the SH3 domain may be another reason for the observed phenomenon; however as the reference surface does not feature any binding, this hypothesis is unlikely. Interestingly, more specific immobilization using α -GST antibodies that are first immobilized on the biosensor surface prior to applying a commercial GST-SH2 domain to the surface also did not cause a different result. However, comparing the rate of dissociation of the proteins from the SPR surface it can be noticed that the non-phosphorylated ER β LBD features a more rapid displacement from the immobilized SH2 domain than the phosphorylated ER β LBD (Figure 25). This effect might indicate a potential influence of phosphorylation on the stability of the ER β – SH2 interaction and needs to be investigated more in detail. In conclusion, the current SPR setup shows that SPR is not a suitable method to detect an interaction between the SH2 domain of c-Src and the ER β LBD via the phosphorylated tyrosine on position 488. Future protein-protein interaction experiments using isothermal titration calorimetry (ITC) may be a more promising method to discover the interaction between the c-Src kinase and the ER β LBD.

4.8 Conclusions

A protein semi-synthesis method to generate correctly folded and active ER LBD proteins was successfully developed. This method allowed, for example, the generation of the ER β LBD with a phosphorylated tyrosine (pY₄₈₈), not accessible via the typical biochemical enzymatic approach. The introduction of a phosphorylation on the tyrosine leads to a higher flexible character of helix 12 and a higher stability of the ER β LBD/E₂/cofactor peptide complex as seen by CD measurements. This is accompanied by a differentiating effect on coactivator peptide recruitment and an increased affinity for coactivator peptides in the absence of ligand and in the presence of the antagonist 4-hydroxytamoxifen, as the most prominent effects. The observed small differences in cofactor binding for the phosphorylated ER β LBD in the presence of E₂ indicate a fine-tuning effect of phosphorylation in terms of cofactor binding. In the absence of ligand, the helix 12 of the phosphorylated LBD could more easily be positioned to generate the cofactor binding surface, AF-2, on the LBD and thus increase peptide binding. The phosphorylation could, for example, stabilize the agonistic conformation by direct interactions with nearby regions in the absence of ligand or in the presence of antagonist. Further, phosphorylation could influence the coactivator binding indirectly by triggering secondary binding interactions of amino acids in and around the LXXLL motif with the (phosphorylated) LBD surface, again contributing to fine-tuning of the selectivity profile of the peptide binding. Interestingly, a ligand-independent interaction with cofactors has been also observed for constitutively active ER β tyrosine mutants in cellular experiments. Replacement of this tyrosine by amino acids with smaller side chains or decreased hydrophobicity resulted in ER α and ER β mutants that display ligand-independent activation equivalent to the level of transcriptional activity observed for the wild-type receptor in the presence of E₂. The ligand independent binding observed in our studies for the semi-synthetic phosphorylated ER β LBD might thus be correlated to the previously observed cellular effects.

For the first time, the structure of a semi-synthetic nuclear receptor featuring a post-translational modification could be successfully crystallized and solved. Interestingly, phosphorylation on tyrosine 488 had no influence on the structural conformation of the E₂-bound LBD. The phosphorylated tyrosine in the agonist bound ER β LBD sticks out of the surface of the protein and thus, is an ideal target for interactions with different proteins as the SH2 domain of c-Src kinase. Access to the NR constructs with control over PTMs provides an entry to study the effects of these PTMs on the regulation of NR protein-protein interactions.

4.9 Experimental Section

General. Unless stated otherwise, all reagents and chemicals were obtained from commercial sources and used without further purification. All other chemical reagents were purchased from Novabiochem, Aldrich-Sigma, Fluka and Acros. All biological reagents were purchased from Invitrogen, Fermentas, New England Biolabs, Qiagen, Novagen, Stratagen, Amersham Bioscience, Serva, Sigma-Aldrich, Thermo Scientific and Pharmacia. LC-ESI-MS was carried out by using an Agilent 1100 series binary pump together with a reversed phase HPLC C18 column (Macherey-Nagel) and a Finnigan Thermoquest LCQ. If not otherwise stated, the following gradient program was used for analytical LC-MS: flow: 1 mL/min, solvent A: 0.1% HCO₂H in H₂O, solvent B: 0.1% HCO₂H in CH₃CN, A/B: 90/10 (0-1 min) to 0/100 (over 10 min). Purification of products by RPHPLC was performed in an Agilent 1100 Series Purification Platform using a NUCLEODUR® C18 Gravity preparative column from Macherey-Nagel (21 x 250 mm) and flow rate of 25 ml/min. The products were eluted by using different solvent gradients of solvents A and B (solvent A = 0.1% TFA/H₂O; solvent B = 0.1% TFA/CH₃CN). UV signal at 210 nm was used for detection. Primers used for all the cloning procedures were supplied by MWG Biotech and all restrictions enzymes were purchased from New England Biolabs or Fermentas.

Peptide synthesis. Fmoc-protected amino acids^[38] were purchased from MultiSyntech and Novabiochem in their appropriately protected form. Solid phase peptide synthesis^[22] was carried out with a Syro II automated peptide synthesizer (MultiSynTech GmbH) utilizing Fmoc-Gly-Wang resin with a loading of 0.78 mmol/g (Novabiochem) and Fmoc-Rink amide MBHA resin (Novabiochem) with a loading of 0.7 mmol/g, respectively and standard Fmoc-protected amino acids; the average batch size was 93.6 mmol. The α -amino Fmoc protecting group was removed using 20% piperidine in DMF. Double amino acid coupling mixtures were prepared by dissolving 4 equivalents of the appropriate amino acid in a mixture of HOBt and addition of a solution of DIC^[39]. After completion the cleavage from resin and the final deprotection was achieved with a mixture of TFA/H₂O/EDT/TIS (94:2.5:2.5:1)^[40] after a reaction time of 2 h. Peptides corresponding to the non-phosphorylated NH₂-CKNVVPVYDLLLEMLNAHVLRG-CO₂H and the phosphorylated ER β LBD NH₂-CKNVVPVpYDLLLEMLNAHVLRG-CO₂H were obtained in 10% and 22% of yield, respectively and a purity of >95%. Peptides for crystallization featuring the sequence NH₂-CKNVVPVpYDLLLEMLNAHVLRG-CO₂H were synthesized by W. Adriaens and were obtained in 8% of yield and a purity of >95% and the Src 1 Box 2 peptide with the sequence NH₂-LTERHKILHRLQEGSPSD-CO₂H was obtained in 28% yield and a purity of >99%. Incorporation of the phosphorylated Y₄₈₈ was achieved using the Fmoc-Tyr-(PO(NMe)₂)-OH building block, which was hydrolyzed to the corresponding phosphate by dissolving the peptide cleaved from the resin in a mixture of TFA/H₂O (1:1) overnight at 4 °C. The peptides were precipitated and washed three times by ice-cold diethylether followed by purification with preparative HPLC on a C18 column using a H₂O-ACN gradient and lyophilization. The purity (>95%) and molecular weights of the final compounds was determined using analytical HPLC.

Polymerase Chain Reaction. PCR^[41] was performed using a T3000 thermocycler (Biometra) or a Mastercycler eppgradient (Eppendorf). The general PCR preparation was consisted of a total volume of 50 μ l containing 1 x *Pfu* DNA polymerase reaction buffer (Fermentas), 5-50 ng dsDNA template, 125 ng of each primer (MWG Biotech) and 200 μ M dNTP mix (Fermentas) and 2.5 U/ μ l *Pfu* DNA polymerase (Fermentas). The amplification of DNA resulted from an initial denaturation at 95 °C for 5 min followed by 30-40 cycles of 30 s denaturation at

95 °C, 45 s of hybridization at specific temperatures (5°C lower than melting temperatures of the primers) and DNA synthesis for 2 min at 72 °C. The last step was a terminal DNA synthesis for 5 min at 72 °C. Control and purification of the amplification product was performed by agarose gel electrophoresis (0.8% agarose) in 40 mM Tris/Acetate pH 7.6, 1 mM EDTA, and SYBR Safe DNA gel stain (Invitrogen). Isolation of DNA from the gel was performed using the Qiaquick PCR purification kit or the Qiaquick gel extraction kit (Qiagen). Isolation of plasmid DNA from *E. coli* was performed using the QIAprep Spin Miniprep Kit (Qiagen).

Cloning, expression and purification.

ERβ LBD with an N-terminal cleavable His₆-tag. pET15-hERβLBD expressing hERβLBD (residues 260-502) with an N-terminal His₆-tag was constructed by amplifying an *NdeI/BamHI* fragment of pET15-hERβ, provided by Peter Donner (Bayer-Schering-Pharma AG) via PCR (mastercycler eppgradient, Eppendorf) using the forward primer 5'-TTTTTTCATATGCTGGACGCCCTGAGCCCCGAGCAG-3' and the reverse primer 5'-TTTTTTGGATCCTCACCCGCGAAGCACGTGGGCATTCAGCATCTC-3' and subsequent subcloning into the vector pET15b (Novagen). The 0.24 kb fragment was digested with *NdeI/BamHI* and ligated with the newly double digested pET15b *E. coli* expression vector using quick T4 DNA Ligase (New England Biolabs). The correct in-frame sequence of the resulting plasmid was confirmed by DNA sequencing (MWG Biotech AG), subsequently transformed into a chemical competent high-density culture of *E. coli* BL21 (DE3) host cells (Stratagene) and grown in 4 L TB medium using selection with ampicillin (100 µg/ml). The cultures were incubated at 37 °C till an OD₆₀₀ of ~1.2 and after cooling down to 18 °C protein expression was induced by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 100 µM. The cells were grown for additional 18-20 h at 15 °C and if necessary 17-β-Estradiol (E₂, Serva) was added to a final concentration of 10 µM. Subsequently cells were harvested by centrifugation (Beckman Coulter, Avanti J26 XP; 4500 rpm, 20 min, 4 °C) and dissolved in a buffer of 1xPBS, 370 mM NaCl, 40 mM Imidazol, pH 8 and 10 % glycerol (10 ml/g of cells) or stored at -80 °C till further use. Subsequently, the bacterial suspension was lysed by up to four cycles under a pressure of 80 bar via shear forces using a micro fluidizer (Microfluidizer 1109) and centrifuged (Beckman Coulter, Avanti J25; 20.000 rpm, 30 min, 4 °C). The soluble cell lysate was immobilized on an equilibrated Nickel-NTA agarose column (HisTrap HP, 5 ml, Amersham Biosciences) using Fast Protein Liquid chromatography (Äkta FPLC, Amersham Biosciences) and washed with lysis buffer to remove all non-bound proteins. The His₆-ERβ LBD fusion protein was finally eluted via an imidazole gradient using elution buffer (1xPBS, 370 mM NaCl, 500 mM Imidazol pH 8 and 10% glycerol). Fractions containing the fusion protein were combined and desalted on a Sephadex G25 PD-10 column (Amersham Biosciences) using desalting buffer (20 mM Tris, 25 mM NaCl, 10 % glycerol and 0.05% β-Octylglycosid). Purity and characterization of the eluted fractions was established by SDS-PAGE using a molecular weight marker (PageRuler Plus Prestained Protein Ladder, Fermentas) and photometric determination of protein concentration using Nanodrop at a wavelength of 280 nm.

ERβ LBD for crystallization. ERβ₂₆₁₋₅₀₀ was amplified from pET15-hERβ by PCR using the forward primer 5'-GAACCATGGACGACGCCCTGAGCCCCGAGCAGCTAGTG-3' and the reverse primer 5'-GGACTCGAGTTAGTCGTCGAAGCACGTGGGCATTCAGCATCTC-3' to introduce the restriction site for *NcoI* and *XhoI*. The primers encode for three extra asp codons, one before the codon for D₂₆₁ and two after L₅₀₀. The expressed LBD thus featured the following sequence: MD[D₂₆₁-L₅₀₀]DD. The PCR fragment was inserted into the *E. coli*

expression vector pET16b (Novagen), double digested with the endonucleases *NcoI* and *XhoI*. The ERβ LBD₂₆₁₋₅₀₀ was overexpressed from high-density culture of *E. coli* BL21 (DE3) host cells (Statagene). Harvested cells were lysed using a micro fluidizer in a buffer containing 20 mM Tris-HCl pH 7.5, 0.5 M NaCl, 5 mM DTT, and 1 mM EDTA (10 ml/g of cells). Clarified lysate was flowed through a pre-equilibrated 3 ml estradiol-Sepharose column (PTI Research, Inc.) and washed with 250 ml of 10 mM Tris-HCl, pH 7.5 containing 0.5 M NaCl and 1 mM EDTA (buffer A). The column was then re-equilibrated with 50 ml of 10 mM Tris-HCl, pH 7.5, 0.2 M NaCl, and 1 mM EDTA (buffer B), and then the protein was carboxymethylated using 50 ml of buffer B containing 5 mM iodoacetic acid. After incubation for 1 h at 4 °C, the column was washed by 400 ml of buffer A, followed by elution in 20 ml buffer A containing 50-200 μM estradiol. Subsequently partial contaminations were removed by size exclusion chromatography (Sephadex 75, HiLoad 26/60, GE Healthcare Biosciences). Fractions containing the purified ERβ LBD₂₆₁₋₅₀₀ were recombined and the buffer was exchanged into buffer B containing 5 mM DTT by passing the solution through a Sephadex G25 PD-10 column (Amersham Biosciences). Finally, the eluted ERβ LBD₂₆₁₋₅₀₀ was concentrated to 10-12 mg/ml using Amicon ultra centrifuge tubes (MWCO 10 kDa) and characterized by SDS-PAGE and photometric determination of protein concentration using Nanodrop at a wavelength of 280 nm.

ERβ LBD α-thioester. For the construction of the ERβ₂₆₀₋₄₈₀-intein-CBD expressing plasmid, the ERβ₂₆₀₋₄₈₀ fragment was amplified by PCR from pET15-hERβ (a kind gift from Dr. Peter Donner, Bayer-Schering Pharma AG) using the forward primer 5'-TTTTTCATATGCTGGACGCCCTGAGCCCCGAGCAG-3' and the reverse primer 5'-TTTTTGTCTTCTGCACTTCATGTTGAGCAGATGTTCCATGCCCTTGTTACT-3', introducing the restriction sites for the endonucleases *NdeI* and *SapI*. This fragment was subsequently inserted into a double digested pTWIN1 *E. coli* expression vector (New England Biolabs). The constructed expression plasmid was transformed into *E. coli* BL21 (DE3) cells and protein expression was performed after cell grows and induction with IPTG. After sedimentation and cell lysis with a micro fluidizer in a buffer of 20 mM HEPES/NaOH (pH 7.3), 500 mM NaCl, 1 mM EDTA, 10% glycerol, 20 μM PMSF, 0.1 mM TCEP, 0.1% Triton-X100 and if necessary 100 μM E₂ the protein extract was centrifuged again. The soluble protein suspension was subsequently applied to a 30 ml pre-equilibrated chitin bead column (New England Biolabs) at 4°C and washed with 10 column volumes of buffer containing 20 mM HEPES/NaOH (pH 7.3), 1 M NaCl, 1 mM EDTA, 0.1 mM TCEP, 10 % glycerol, 1mM ATP, 2mM Mg₂Cl and if necessary 10 μM E₂. To remove ATP, MgCl and the excess of NaCl the column was washed with two column volumes of buffer containing 20 mM HEPES/NaOH (pH 7.3), 500 mM NaCl, 1 mM EDTA, 0.1 mM TCEP, 10 % glycerol and if necessary 10 μM E₂. On-column cleavage of the intein-tag was induced by equilibrating the chitin beads with two volumes cleavage buffer (20 mM HEPES/NaOH (pH 8.5), 500 mM NaCl, 1 mM EDTA, 10 % glycerol, 200 mM MESNA and if necessary 10 μM E₂) for 24 h at 4 °C. Elution of the protein was performed by washing the column with buffer cleavage buffer. Elution fractions were collected and pooled after which the cleavage step was repeated to gain more thioester terminated proteins. The proteins were concentrated using Amicon ultra centrifuge tubes and the efficiency of intein cleavage was confirmed by SDS gel electrophoresis and ESI mass spectroscopy.

ERβ LBD α-thioester with N-terminal His₈-tag. pTWIN1-HishERβLBD expressing hERβLBD (residues 260-480) with a N-terminal His₈-tag and a C-terminal intein tag followed by a chitin binding domain (CBD) was constructed by subcloning an *AflIII/SapI* fragment of pET15-hERβLBD into the plasmid pTWIN1-His₈ digested

with *NcoI/SapI*. Used primers were: 5'-GGCCATACATGTCTGGTCTGGACGCCCTGAGCCCCGA-3' and 5'-TTTTTGTCTTCTGCACTTCATGTTGAGCAGATGTTCCATGCCCTTGTTACTCGC-3'. The pTWIN1 - His₈ vector was created by cloning a His₈-tag into the vector pTwin1 using the primers 5'-TATGGAAGCGAG-CCACCATCACCATCACCATCACCATG-3' and 5'-CATGGCATGGTGTATGGTGTATGGTGTATGGTGGC-TCGCTTCCA-3' and ligating the PCR fragment with the *NdeI/NcoI* digested pTwin1 vector. Expression and purification of the ERβ LBD α-thioester with N-terminal His₈-tag was performed like described for the ERβ LBD α-thioester without His₈-tag via intein thiolysis.

ERβ LBD α-thioester for crystallization. ERβ₂₆₀₋₄₈₀ was amplified from the vector pET15-hERβLBD by PCR using the forward primer 5'-TTTTTTCATATGGACGACGCCCTGAGCCCCGAGCAG-3' and the reverse primer 5'-TTTTTGTCTTCTGCACTTCATGTTGAGCAGATGTTCCATGCCCTTGTTACTCGC-3' to introduce the restriction sites for *NdeI* and *SapI*. The PCR fragment was inserted into the double digested vector pTWIN1 by ligation. The primers encode one extra asp codon before the codon for D₂₆₁. The expressed LBD thus has the following sequence: MD[D₂₆₁-K₄₈₀]. Expression of the ERβ LBD α-thioester for crystallization was performed in *E. coli* as described before.

Cleavage of a His₆-tag. If necessary enzymatic cleavage of the His₆-tag was performed using a thrombin protease (1 unit/μl; Amersham Biosciences). One cleavage unit was incubated with 100 μg of fusion protein in desalting buffer at room temperature over night. After cleavage the thrombin was removed from the sample by addition of Benzamidine sepharose beads (GE Healthcare), which bind thrombin. The beads were three times washed in distilled water before usage. The beads (1 μl per 1 μl protein) were incubated with the sample for about 30 minutes at 4 °C under slow stirring followed by a 10 minutes centrifugation step at 4500 rpm to remove the sepharose.

SH2/SH3 domain: The SH2/SH3 domain (aa 84 - 248) of the c-Src kinase was generated as fusion protein with an N-terminal His₆-tag. The template DNA for the PCR was drawn from the human v-Src sarcoma viral oncogene homolog (Ultimate ORF clone IOH 12563, Invitrogen). Used forward primer: 5'-AAGTTCTGTTTC-AGGGTGGTGGAGTGACCACCTTTGTGG-3' and used reverse primer: 5'-CTGGTCTAGAAAGCTTCA-GCACACGGTGGTGGAGGCG-3'. The PCR product was gel purified and mixed with the linearized pTriEX-2-based^[37] expression vector pOPINF (Dortmund Protein Facility, DPF), allowing a ligation-free *in vivo* cloning strategy^[42]. After transformation of the DNA mix into *E. coli* One Shot OmniMax cells (Invitrogen) the correct open reading frame was confirmed by DNA sequencing (MXG Biotech AG). Expression of the His-SH2/SH3 domain was performed in *E. coli* BL21 (DE3) Codon + RIL expression cells (Stratagene) overnight at 15 °C using 100 μM IPTG. Upon cell disruption, the soluble protein extract was applied to a Ni²⁺ NTA column (His Trap FF crude 5 ml, Amersham Biosciences) using the Äkta FPLC. The tagged target protein binds to the matrix and unbound protein is washed out with a buffer of 50 mM Na₂HPO₄, 300 mM NaCl, 1 mM TCEP and 20 mM imidazol pH 8.0. For cleavage of the His-tag, the column is equilibrated for 5 h with cleavage buffer (50 mM Hepes, 150 mM NaCl and 1 mM TCEP pH 8.0) containing 0.7 CV PreScission Protease (GE Healthcare). Purity of the eluted SH2/SH3 domain, cleaved from the His-tag, was confirmed by SDS-PAGE.

SH2 domain: pOPINF-SH2 expressing the SH2 domain of c-Src (aa 151 – 248) was constructed by amplifying a fragment of the human v-Src sarcoma viral oncogene homolog (Ultimate ORF clone IOH 12563, Invitrogen) via PCR using the forward primer 5'-AAGTTCTGTTTCAGGGTTGGTATTTTGGCAAGATCACCAGACG-3' and the reverse primer 5'-CTGGTCTAGAAAGCTTCAGCACACGGTGGTGAGGCG-3'. Expression and purification of the SH2 domain was performed as described above.

Size exclusion chromatography of the ERβ LBD. Size exclusion chromatography was used for preparative purification and analytical confirmation of the purity of proteins. The FPLC (Åkta FPLC, GE Healthcare, Buchler) was performed with filtrated and degassed buffers. For preparative approaches a HiLoad 26/60 Sephadex 75 column was used while analytical experiments were carried out using a HiLoad 16/60 Sephadex 75 column. The columns were equilibrated with two volumes of target buffer at 4 °C. The protein was concentrated to more than 10 mg/ml in a volume of ~1.5 ml and applied to the column with an approximately flow rate of 1 ml/min. The eluat was collected in 2 ml fractions and analyzed using SDS-PAGE.

Expressed protein ligation of ERβ LBD₂₆₀₋₄₈₀ α-thioester to ERβ LBD₂₆₀₋₅₀₂. Native chemical ligation of the 200 μl ERβ LBD α-thioester (200 μM) with 200 μl peptide (4 mM) was performed in a buffer containing 20 mM HEPES/NaOH (pH 7.3), 500 mM NaCl, 1 mM EDTA, 10 % glycerol and if necessary 10 μM E₂ in the presence of 200 μl 150 mM MESNA (Sigma-Aldrich). The reaction was incubated for 48 h at 4 °C with slight agitation. Reaction mixtures containing a protein construct bearing a His₈-tag on the N-terminus were passed through a Ni-NTA column (HisTrap HP, 1 ml, Amersham Biosciences) to remove unligated peptide. Proteins without His₈-tag were extensively dialyzed for three days at 4°C against ligation buffer without MESNA using Slide-A-Lyzer Dialysis Cassettes (20 MWCO, Thermo Scientific) or were applied to a size exclusion column to remove the excess of peptide. Efficiency of the ligation reaction was determined by SDS gel electrophoreses and ESI mass spectroscopy.

Enzymatic phosphorylation of tyrosine 488. The enzymatic phosphorylation was performed by E. van der Vaart (Organon). The ERβ LBD (~115 nmol) was incubated for 1 h at 30°C with two units of a c-Src kinase (Upstate; MW: 61.7 kDa) and 100 μM ATP, in kinase buffer of pH 7.2 (10 mM Tris-HCl, 10 mM MgCl₂, 0.01% Tween-20, 0.05% NaN₃, 1 mM DTT and 2 mM MnCl₂). The reaction was stopped by addition of SDS-sample buffer (0.125 M Tris, 20% (v/v) glycerol, 4% (w/v) SDS, 0.01% (w/v) bromophenol blue and 10% (v/v) β-mercaptoethanol) in a 1:1 ratio and subsequent SDS-PAGE.

Immunodetection of the phosphorylated tyrosine. For the detection of a successful phosphorylation on Y₄₈₈ of the ERβ LBD, the protein was gel electrophoretic separated by SDS-PAGE and subsequently immobilized on a nitrocellulose membrane (Schleicher & Schuell)^[43] using an semi-dry transfer cell (BioRad) with buffer containing 20 mM Tris, 150 mM glycine, 20% (v/v) methanol and 0.1% SDS. The transfer was performed for 30 min and 10 V at room temperature. For immunodetection of specific immobilized proteins first the nitrocellulose was incubated in blocking buffer (50 mM Tris/HCl, pH 7.8, 150 mM NaCl, 1 mM MgCl₂, 2% (w/v) milk powder) over night at 4 °C. Subsequently the nitrocellulose was incubated for 1 h at 4 °C with a α-phosphotyrosine antibody conjugated with FITC (Sigma-Aldrich) in a 1:1000 dilution. Unspecific bound antibody was removed by three washing steps each for 10 min with TBST (100 mM Tris/HCl, pH 8.0, 150 mM

NaCl, 1 mM MgCl₂ and 0.1% Tween 20). Binding of a secondary immunopure Anti-mouse IgG (H+L) – Horseradish Peroxidase antibody (Pierce) in a 1:15000 dilution in blocking buffer for 1 h at 4 °C. After washing visualization of tyrosine phosphorylation was performed via incubating the blot in a 1:1 mixture of staining solutions (SuperSignal West Pico Luminol Enhancer Solution and SuperSignal West Pico Stable Peroxide Solution, Thermo Scientific) for 5 min. Subsequently the blot was covered with an X-ray radiograph (CL-XPosure Film, 5x7 inch, Thermo Scientific) for one minute in the dark followed by the development of the radiograph. Alternatively to the luminol substrate method, visualization was performed using a DAB solution containing H₂O₂, which was converted to colored insoluble precipitate.

Circular Dichroism Spectroscopy. For the CD measurements the stock solution of the ERβ LBD was buffer-exchanged into 20 mM NaH₂PO₄ buffer (pH 7.7) using a Sephadex G25 PD-10 column (Amersham Biosciences). CD spectra were recorded on a Jasco J-815 CD spectrometer equipped with a peltier-type temperature controller (PTC-423S) at 10 °C using a 1 mm path length quartz cell (0.5 nm data pitch, continuous scanning mode, 100 nm x min⁻¹ scanning speed, 0.5 nm band width) at a protein concentration of 4 μM. The measured wavelength range was from 190 nm to 250 nm. Ten scans were collected and averaged. Mean residue ellipticity [θ] was calculated using the following equation:

$$[\theta]_R = \frac{[\theta]}{10 \times (n-1) \times c \times l} \left[\frac{\text{deg}^{-3} \times L}{\text{mol} \times \text{cm}} \right]$$

where n is the number of amide bonds, c is the concentration (mol/L), l is the pathlength (cm), $[\theta]$ is the measured ellipticity (deg⁻³) and $[\theta]_R$ is the mean residue ellipticity. For thermal stability measurements of the ERβ LBD, the ellipticity at 222 nm was followed as a function of temperature in 1 cm cuvettes with a volume of 600 μl. Points were taken from 40 to 80 °C at 1°C/min intervals. E₂ was added in a final concentration of 10 μM and Src 1 Box 2 peptide was used at a final concentration of 40 μM. To determine the apparent T_M and enthalpy of folding, data were fit to the following equations, assuming that all the ellipticity change was due to a two-state transition between the folded and unfolded receptor:

$$k = \exp\left\{\left[\frac{\Delta H}{RT}\right]\left[\left(\frac{T}{T_{Mobs}}\right) - 1\right]\right\}$$

$$y = k \div (1 + k)$$

$$\theta_{obs} = (\theta_{max} - \theta_{min})y + \theta_{min}$$

where θ_{obs} is the ellipticity found at any temperature, θ_{max} is the maximum ellipticity corresponding to fully folded ERβ LBD, θ_{min} is the ellipticity corresponding to the unfolded ERβ LBD, ΔH is the apparent van't Hoff enthalpy of folding, T_{Mobs} is the midpoint of the folding transition, and R is the gas constant. θ_{max} , θ_{min} , T_{Mobs} , and ΔH were estimated by non-linear least-squares curve fitting using the Spectra analysis program of the Jasco Manager version 2 and choosing the denaturation analysis option. Note that the enthalpy calculated by this method is not a true thermodynamic parameter, because the unfolding was not reversible. In addition, no corrections were made to take into account the stoichiometry of the receptor – ligand – peptide complexes. The enthalpy measurements are given only as estimates of the cooperativity of the unfolding transition for comparative purposes.

NR-coregulator interaction profiling. The On-Chip assay was performed in the absence and presence of ligand. For the preparation of the reaction mixtures purified His-ER β LBD was thawed on ice. TR-FRET Coregulator buffer E (Invitrogen, Carlsbad, USA) was complemented with DTT to a final concentration of 5 mM and all reagents were dissolved in this buffer. The final reaction mixture comprised 25 nM Alexa488-conjugated α -His-antibody (Invitrogen, Carlsbad, USA) and 4 nM His-ER β LBD and if necessary 0.01 mM 17- β -estradiol (Serva) or *trans*-4-hydroxytamoxifen (Sigma-Aldrich). Prior use the ligand was dissolved in DMSO and diluted to a final DMSO concentration of 0.1%. Samples without ligand were also complemented with 0.1% DMSO for comparison. To confirm the sensitivity and reproducibility of the assay dilutions of the ER β proteins were performed. The dilution concentrations were 1/10, 1/50, and 1/100. All assays were performed in a PamStation®-4 controlled by EvolveHT software (PamGene International BV, 's-Hertogenbosch, The Netherlands) at 20°C, at a rate of 2 cycles per minute. Nuclear Receptor PamChip Arrays (PamGene International BV, 's-Hertogenbosch, The Netherlands) contained 53 peptides on each spot. The arrays are made of a porous metaloxide carrier to which the peptides are spotted by means of Piezzo technology. Each spot has a diameter of 100 μ m and due to the porous structure the surface area is ~500 times larger than calculated based on spot diameter. A spot contains ~106 pores, each with a diameter of 0.2 μ m and a length of 60 μ m. In addition, pores are branched and interconnected. Arrays were incubated for 20 pump cycles with 25 μ l blocking (buffer 1% BSA, 0.01%, Tween-20 in Tris-buffered Saline) and then aspirated. Each array, was incubated with 25 μ l of a solution of His-ER β LBD, fluorescent α -His-antibody and if necessary ligand. Unspecific binding and antibody effects were taken into account by subtracting the simultaneous response from a reference array containing no His-ER β LBD. The solution was pumped through the porous peptide-containing membrane for 81 cycles at a rate of 2 cycles per minute and a image of every array was obtained every 20 cycles by a CCD camera based optical system integrated in the PamStation-4 instrument. For imaging a camera filter for the FITC dye with an excitation wavelength of 475 nm and an emission wavelength of 535 nm was used. Peptide microarray data analysis consisting of automated spot finding and quantitation, followed by calculation of binding velocities was performed by Bionavigator software (PamGene International BV, 's-Hertogenbosch, The Netherlands).

Cofactor recruitment FRET assay. The TR-FRET assay was performed with 31 peptides in the absence and presence of ligand (E₂ and HT). The optimized reaction mixture contained 10 nM of purified His-ER β LBD, 1.25 nM europium-labelled α -His antibody (Eu; Perkin Elmer Life Science), 80 nM allophycocyanin-labelled Streptavidin (SA-APC; Perkin Elmer Life Science), if necessary 10 μ M E₂ and increasing concentrations of biotinylated cofactor peptide, in the assay buffer (50 mM Tris-HCl, pH 7.4, 50 mM KCl, 1 mM EDTA, 0.1 mg/ml BSA, 1 mM DTT). The dilutions of peptide ranged from 1 μ M to 0.01 nM. Each experiment was performed in duplicate in 384-well plates (Packard Optiplat) by adding 20 μ l of a receptor/ Eu-labelled α -His-antibody mixture to 10 μ l of a peptide/ SA-APC mixture followed by adding 20 μ l of 10 μ M E₂. The reactions were mixed, centrifuged and routinely incubated overnight at 4 °C. Fluorescence at 665 nm was measured using an EnVision (2102 multilabel) Counter with a setting of 50 μ sec time delay, excitation at 340 nm, and emission at 665 nm. All data were means of two separate experiments. To directly visualize the effect of a ligand on recruitment of the different peptides the data were used for sigmoid curve plotting using Origin 7.5 (Scientific Graphing and Analysis Software, OriginLab Corp.) and the apparent EC₅₀ values of peptide binding were determined.

Co-crystallization of the ER β LBD in complex with cofactor peptide and estradiol. For crystallization, both the γ -phosphorylated and the non-phosphorylated ER β LBD/E₂ complex was concentrated to 12 mg/ml in a buffer containing 0.2 M NaCl, 1 mM EDTA, 5 mM DTT and 10 mM Tris-HCl at pH 7.5 and then mixed with the SRC-1 Box 2 peptide at a molecular ratio of 1.5 : 1 peptide to protein – ligand complex. Prior to that the peptide was dissolved in crystallization buffer to a concentration of 8.97 mM and extensively dialyzed against the same buffer for three days at 4 °C using a dialysis membrane from Spectra/Por (MWCO: 1 kDa). In order to crystallize the ER β LBD complex, initial screenings employing JCSG+, JCSG Core I, JCSG Core II, JCSG Core III and JCSG Core IV from Qiagen were performed at 20 and 4 °C using the sitting drop vapor diffusion. The protein concentration in the setups was 12 mg/ml; 0.1 μ l protein solutions were automatically mixed with 0.1 μ l reservoir solution in 96 well Corning pZero plates using a phoenix pipetting robot. The sitting drops were equilibrated against reservoirs with a volume of 70 μ l. After one to two days, hexagonal crystals appeared in several conditions typically containing 0.2 M of a diverse salt and 20% PEG 3350 and reaching a maximal size of 220 x 88 x 73 μ m within three days. Initial screening experiments with the phosphorylated ER β LBD complex lead to crystals with the same size and shape growing in similar conditions. Reproduction of these crystals of both ER β LBD complexes were performed in 24 well EasyXtal DG-tools (Qiagen) using the hanging drop vapor diffusion method. Drops with a size of 2 to 5 μ l using a different reservoir to protein ratio were manually mixed at 4 °C and equilibrated against reservoirs with a volume of 1 ml. Optimal crystals were grown over night in 4.5 μ l drops with a protein solution to reservoir ratio of 3 : 1.5 in a condition containing 0.2 M Ammonium Acetate, 0.1 M BIS-TRIS pH 5.5 and 17% PEG10000 (v/v) at 4 °C. These crystals reached a final size of 0.5 x 0.4 x 0.3 mm within five days and were of fine quality as judged by light microscopy. After soaking in reservoir solution supplemented with glycerol to 20%, the crystals were cryo-cooled in liquid nitrogen for X-ray data collection at 100 K. Since the ER β LBD has been crystallized before and suitable protein structures could be found in the Protein Data Bank, the crystal structure could be solved by molecular replacement of a ER β LBD/ligand/peptide complex (PDB code: 1u9e)^[31b] using the program PHASER^[34].

A native data set was collected from an optimized ER β LBD complex crystal at 100 K on a Nonius AXS MICRO star at a wavelength of 1.548 Å using a MAR dtb detector. The crystal-to-detector distance was 150 mm and the oscillation range was 1.0°. Data were indexed, integrated and scaled with the XDS package^[33]. The crystals diffracted to a maximal resolution of about 2.2 Å and a check for possible systematic absences revealed that the crystals belonged to space group $P3_1$, with one homo dimer per asymmetric unit. According to Matthews^[44], protein crystals typically exhibit a solvent content of 30 to 70%. Using the information about the volume of the unit cell and the molecular weight of the used protein complex, the number of molecules per asymmetric unit (ASU) can be estimated. The calculation of the Matthews parameters for a molecular weight of 29533 Da per ER β LBD/E₂/peptide complex and a cell volume of 475250 Å³, probably resulted in two complexes with a $V_M = 2.68$ Å³/Da and a solvent content of 54.17% in an asymmetric unit.

In order to improve the maximal resolution of the phosphorylated ER β LBD complex an additional data set was collected at 100 K on the *Swiss Light Source* (SLS, Villigen, Switzerland) beamline X10SA (PXII) using a MAR 225 CCD detector. The crystal-to-detector distances were 140 mm and the oscillation range for the data set was 1°. These data were also indexed, integrated and scaled with the XDS package^[33]. Data set statistics of both crystal structures are given in Table 5. The structure building was performed by calculation of the electron density maps using the program Coot^[45]. The iterative structure refinement was carried out with the program

REFMAC^[46] from the program package CCP4 suite^[47]. All structural representations were prepared with pymol (www.pymol.org)

Surface Plasmon Resonance: All experiments were performed on a Biacore T100 (GE Healthcare) on CM5 chips. The running buffer used was HBS buffer (10 mM HEPES, 150 mM NaCl, 3 mM EDTA) supplemented with 0.005% surfactant P20 (Roche) at pH 7.4. For functionalizing the SH2 domain was dissolved in Acetat buffer at pH 4.5 to a final concentration of 50 µg/ml. Immobilization was performed at 25 °C using the standard EDS/NHS protocol. Subsequently, the surface was flowed with increasing concentrations of ERβ LBD dissolved in HBS buffer with P20 at a flow rate of 30 µl/min. Regeneration after each run was performed using buffer containing 10 mM glycine-HCl at pH 2.2 and 0.05% SDS. Unspecific binding and buffer effects were taken into account by subtracting the simultaneous response from a reference surface containing no SH2 domain, but functionalized with ethanolamine instead. In case of α-GST antibody (Biacore) immobilization of the membrane was incubated with a recombinant GST- Src SH2 domain fusion protein (Marligen Biosciences).

List of Cofactor peptides used for the On-Chip studies

Cofactor peptide	Sequence
CBP_LXXLL70_57_80	GNLVDPDAASKHKQLSELLRGGSGS
CBP_LXXLL358_345_368	TADPEKRKLIQQQLVLLLHAHKCQ
CBP_LXXLL358_345_368_C367S	TADPEKRKLIQQQLVLLLHAHKSQ
CBP_LXXLL2067_2055_2077	SVQPPRSISPSALQDLLRRTLKSP
DAX1_LxxML13_1_23	MAGENHQWQGSILYNMLMSAKQT
DAX1_LXXLL146_136_159	GEDHPRQGSILYSLLTSSKQTHV
DAX1_LxxML80_68_90_C69S	FSGKDHPROGSILYSMLTSAKQT
EP300_LXXLL2051_2039_2061	SPLKPGTVSQQALQNLRLTLRSP
EP300_LXXLL81_69_91	GMVQDAASKHKQLSELLRSGSSP
IKBB_LXXLL289_277_299	PLGSAMLRPNPILARLLRAHGAP
IKBB_LXXLL74_62_84	LHLAVIHQHEPFLDFLLGFSAGT
JMJ1C_LXXLL37_25_47	PLVSONNEQGSTLRDLLTTTAGK
KIF11_LXXLL845_833_855_C855S	WVSSLNEREQELHNLLEVVSQCS
NCOA1_LXXLL1435_1421_1441	TSGPQTPQAQQKSLQQLLTT
NCOA1_LXXLL633_620_643	SDGDSKYSQTSHKLVQLLTTTAEQ
NCOA1_LXXLL690_677_700	PSSHSSLTERHKILHRLLEQEGSP
NCOA1_LXXLL749_737_759	ASKKKESKDHQLRYLLDKDEK
NCOA2_LXXLL641_628_651	QSRHLHDSKGQTKLLQLLTTKSDQ
NCOA2_LXXLL690_677_700	STHGTSLKKEKHILHRLLDQSSSP
NCOA2_LXXLL745_733_755	EPVSPKKKENALLRYLLDKDDTK
NCOA2_LXXLL878_866_888	SQSTFNNPRPGQLGRLLPNQNL
NCOA3_LXXLL113_102_123_N-KKK	STGQGVIDKDSLGLPLLQALDG
NCOA3_LXXLL621_609_631	QRGPLESKGHKKLLQLLTCSSDD
NCOA3_LXXLL621_609_631_C627S	QRGPLESKGHKKLLQLLTTSSDD
NCOA3_LXXLL685_673_695	MHGSLLEQKHRIHLKLLQNGNSP
NCOA3_MOUSE_LXXLL1041_1029_1051	HGSQNRPLLRNSLDDLLGPPSNA
NCOA6_LXXLL1491_1479_1501	LVSPAMREAPTSLSQLLDNSGAP
NCOA6_LXXLL887_875_897	PVNKDVTLTSPLLVNLQSDISA
NCOR1_LxxHI2051_2039_2061	MGQVPRTHRLLITLADHICQIITQ
NCOR1_LxxHI2051_2039_2061_C2056S	MGQVPRTHRLLITLADHISQITQ
NCOR1_LxxII2263_2251_2273	GHSFADPASNLGLEDIIRKALMG
NCOR2_LxxHI2135_2123_2145	APGVKQGHQRVVTLAQHISEVITQ
NCOR2_LxxII2342_2330_2352	QAVQEHASTNMGLEAIIRKALMG
NRIP1_LXXLL133_121_143	DSVPKQKQDSTLLASLLQSFSSR
NRIP1_LXXLL185_173_195	KDLRCYGVASSHLKTLKSKSVK
NRIP1_LXXLL185_173_195_C177S	KDLRSYGVASSHLKTLKSKSVK
NRIP1_LXXLL380_368_390	RNNIKQAANNSLLHLHLLKSKQITP

NRIP1_LXXLL500_488_510	KNSKLNHQKVTLLQLLGHKNE
NRIP1_LXXLL713_701_723	EIENLLERRTVLQLLGNPNKG
NRIP1_LXXLL819_805_831	PVSPQDFSFSGKNGLLSRLLRQNDQSYL
NRIP1_LXXLL936_924_946	RSWARESKSFNVLKQLLLENCV
NRIP1_LXXLL936_924_946_C945S	RSWARESKSFNVLKQLLLENSV
PCAF_LXXLL190_178_200	EEDADTKQVYFYLKLLRKSILQ
PPRB_LXXLL604_591_614	HGEDFSKVSQNPILTSLLQITGNG
PPRB_LXXLL645_632_655	VSSMAGNTKNHPMLMNLKDNPAQ
PRGC1_LXXLL144_130_155	DGTPPPQEAEEPSLLKKLLAPANTQ
PRGC1_LXXLL144_134_154	PPQEAEEPSLLKKLLAPANT
PRGC2_LXXLL156_146_166	PAPEVDELSLLQKLLATSYP
PRGC2_LXXLL343_338_358	AEFSILRELLAQDVLCDVSKP
SHP_LxxIL118_106_128	TFEVAEAPVPSILKKILLEEPS
SHP_LXXLL21_9_31_C9S_C11S	SPSQGAASRPAILYALLSSSLKA
TRIP4_LXXLL161_149_171_C171S	FVNLYTREGQDRLAVLLPGRHPS
ZNHI3_LXXLL101_89_111	LQNLKNLGESATLRSLLNPHLR

List of Cofactor peptide used for the FRET studies

Cofactor peptide	Sequence
SCR2_1	Biotinyl-DGQSRLHDSKGQTKLLQLLTTKSDQ
SCR1_2	Biotinyl-CPSSHSSLTERHKILHRLLEQEGSPS
SCR2_3	Biotinyl-QEPVSPKKKENALLRYLLDKDDTKD
SHP_1	Biotinyl-CPCQGAASRPAILYALLSSSLKAVP
SHP_2	Biotinyl-TFEVAEAPVPSILKILLEEPS
D22	Biotinyl-LPYEGSLLLKLLRAPVEEV
D47	Biotinyl-HVYQHPLLSLLSSEHESG
DAX3	Biotinyl-GEDHPRQGSILYSLLTSSKQTHVA
RIP140_5	Biotinyl-LERNNIKQAANNSLLHLLKSQTIP
RIP140_6	Biotinyl-SKNSKLNHQKVTLLQLLGHKNEE
RIP140_8	Biotinyl-PVSPQDFSFSGKNGLLSRLLRQNDQSYL
RIP140_9	Biotinyl-EHRWARESKSFNVLKQLLLENCV
CBP_1	Biotinyl-NLVPDAASKHKQSELLRGGSGS
PGC_1	Biotinyl-DGTPPPQEAEEPSLLKKLLAPANTQ
RIP140_3	Biotinyl-EKDLRCYGVASSHLKTLKKSKVKD
HRCoA2	Biotinyl-AEDRAGRPLPCPSLCELLASTAVK
TRAP220_1	Biotinyl-GHGEDFSKVSQNPILTSLLQITGNG
TRAP220_2	Biotinyl-PVSSMAGNTKNHPMLMNLKDNPAQ
SRC3_1	Biotinyl-SNMHGSLLEQKHRIHLKLLQNGNSP
SRC3_1	Biotinyl-QEQLSPKKKENALLRYLLDRDDPS
SRC1a_4	Biotinyl-TSGPQTPQAQQKSLQQLLITE
SRC1_3	Biotinyl-LDASKKESKDHQLLRYLLDKDEKD
RIP140_1	Biotinyl-MVDSVRKKGQDSTLLASLLQSFSSR
RIP140_7	Biotinyl-SGSEIENLLERRTVLQLLGNPTKG
D30	Biotinyl-HPTHSRLWELLMEATPTM
TIP60	Biotinyl-TLSEDIVDGERAMLKRLLRIDSKC
PERC-1	Biotinyl-KPSAPAPEVDELSLLQKLLATSYP
PERC-2	Biotinyl-SRHHSKASWAEFSILRELLAQDVLCD
EA2	Biotinyl-SSKGVLWRMLAEPVSR

4. 10 References

- [1] a)D. A. Lannigan, *Steroids* **2003**, *68*, 1; b)C. Rochette-Egly, *Cell Signal* **2003**, *15*, 355; c)H. Faus, B. Haendler, *Biomedicine & Pharmacotherapy* **2006**, *60*, 520; d)N. L. Weigel, N. L. Moore, *Mol Endocrinol* **2007**, *21*, 2311; e)N. L. Weigel, N. L. Moore, *Molecular Endocrinology* **2007**, *21*, 2311.
- [2] a)A. R. de Lera, W. Bourguet, L. Altucci, H. Gronemeyer, *Nature Reviews Drug Discovery* **2007**, *6*, 811; b)A. R. de Lera, W. Bourguet, L. Altucci, H. Gronemeyer, *Nat Rev Drug Discov* **2007**, *6*, 811.
- [3] J. M. Wurtz, W. Bourguet, J. P. Renaud, V. Vivat, P. Chambon, D. Moras, H. Gronemeyer, *Nature Structural Biology* **1996**, *3*, 87.

- [4] C. Kremoser, M. Albers, T. P. Burris, U. Deuschle, M. Koegl, *Drug Discovery Today* **2007**, *12*, 860.
- [5] a)D. Schwarzer, P. A. Cole, *Curr Opin Chem Biol* **2005**, *9*, 561; b)V. Muralidharan, T. W. Muir, *Nat Methods* **2006**, *3*, 429; c)T. W. Muir, D. Sondhi, P. A. Cole, *Proc Natl Acad Sci U S A* **1998**, *95*, 6705.
- [6] a)P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. Kent, *Science* **1994**, *266*, 776; b)T. M. Hackeng, J. H. Griffin, P. E. Dawson, *Proceedings of the National Academy of Sciences of the United States of America* **1999**, *96*, 10068.
- [7] a)R. David, M. P. Richter, A. G. Beck-Sickinger, *Eur J Biochem* **2004**, *271*, 663; b)J. P. Pellois, T. W. Muir, *Curr Opin Chem Biol* **2006**, *10*, 487; c)M. P. O. Richter, A. G. Beck-Sickinger, *Protein and Peptide Letters* **2005**, *12*, 777.
- [8] a)G. Castoria, A. Migliaccio, S. Green, M. Di Domenico, P. Chambon, F. Auricchio, *Biochemistry* **1993**, *32*, 1740; b)S. F. Arnold, D. P. Vorojeikina, A. C. Notides, *Journal of Biological Chemistry* **1995**, *270*, 30205; c)F. Barletta, C. W. Wong, C. McNally, B. S. Komm, B. Katzenellenbogen, B. J. Cheskis, *Mol Endocrinol* **2004**, *18*, 1096; d)V. Giguere, A. Tremblay, G. B. Tremblay, *Steroids* **1998**, *63*, 335.
- [9] G. B. Tremblay, A. Tremblay, F. Labrie, V. Giguere, *Cancer Res* **1998**, *58*, 877.
- [10] a)A. Migliaccio, G. Castoria, M. Di Domenico, A. de Falco, A. Bilancio, M. Lombardi, M. V. Barone, D. Ametrano, M. S. Zannini, C. Abbondanza, F. Auricchio, *EMBO J* **2000**, *19*, 5406; b)R. Song, H. I. Hall, R. Frey, *Stat Med* **2005**, *24*, 453; c)L. Varricchio, A. Migliaccio, G. Castoria, H. Yamaguchi, A. de Falco, M. Di Domenico, P. Giovannelli, W. Farrar, E. Appella, F. Auricchio, *Mol Cancer Res* **2007**, *5*, 1213; d)F. Auricchio, A. Migliaccio, G. Castoria, *Steroids* **2008**, *73*, 880; e)B. J. Cheskis, J. Greger, N. Cooch, C. McNally, S. McLarney, H. S. Lam, S. Rutledge, B. Mekonnen, D. Hauze, S. Nagpal, L. P. Freedman, *Steroids* **2008**, *73*, 901; f)S. B. Kim, Y. Umezawa, K. A. Kanno, H. Tao, *ACS Chem Biol* **2008**, *3*, 359.
- [11] L. Zhong, D. F. Skafar, *Biochemistry* **2002**, *41*, 4209.
- [12] W. Bourguet, P. Germain, H. Gronemeyer, *Trends in Pharmacological Sciences* **2000**, *21*, 381.
- [13] S. F. Arnold, J. D. Obourn, H. Jaffe, A. C. Notides, *Molecular Endocrinology* **1995**, *9*, 24.
- [14] U. K. Laemmli, *Nature* **1970**, *227*, 680.
- [15] A. C. U. Steinmetz, J. P. Renaud, D. Moras, *Annual Review of Biophysics and Biomolecular Structure* **2001**, *30*, 329.
- [16] a)P. M. Kane, C. T. Yamashiro, D. F. Wolczyk, N. Neff, M. Goebel, T. H. Stevens, *Science* **1990**, *250*, 651; b)F. B. Perler, E. O. Davis, G. E. Dean, F. S. Gimble, W. E. Jack, N. Neff, C. J. Noren, J. Thorner, M. Belfort, *Nucleic Acids Res* **1994**, *22*, 1125; c)F. B. Perler, *Cell* **1998**, *92*, 1; d)C. J. Noren, J. M. Wang, F. B. Perler, *Angewandte Chemie-International Edition* **2000**, *39*, 450.
- [17] a)M. Q. Xu, F. B. Perler, *EMBO J* **1996**, *15*, 5146; b)S. Chong, F. B. Mersha, D. G. Comb, M. E. Scott, D. Landry, L. M. Vence, F. B. Perler, J. Benner, R. B. Kucera, C. A. Hirvonen, J. J. Pelletier, H. Paulus, M. Q. Xu, *Gene* **1997**, *192*, 271; c)T. C. Evans, Jr., J. Benner, M. Q. Xu, *Journal of Biological Chemistry* **1999**, *274*, 3923; d)S. Mathys, T. C. Evans, I. C. Chute, H. Wu, S. Chong, J. Benner, X. Q. Liu, M. Q. Xu, *Gene* **1999**, *231*, 1.
- [18] a)M. Q. Xu, T. C. Evans, Jr., *Methods Mol Biol* **2003**, *205*, 43; b)M. Q. Xu, T. C. Evans, Jr., *Methods* **2001**, *24*, 257.
- [19] T. Watanabe, Y. Ito, T. Yamada, M. Hashimoto, S. Sekine, H. Tanaka, *J Bacteriol* **1994**, *176*, 4465.
- [20] T. C. Evans, Jr., J. Benner, M. Q. Xu, *Protein Sci* **1998**, *7*, 2256.
- [21] a)J. Fenn, *J Biomol Tech* **2002**, *13*, 101; b)M. Dole, L. L. Mack, R. L. Hines, *Journal of Chemical Physics* **1968**, *49*, 2240.
- [22] R. B. Merrifield, *Journal of the American Chemical Society* **1963**, *85*, 2149.
- [23] T. R. Burke, K. Lee, *Accounts of Chemical Research* **2003**, *36*, 426.
- [24] a)Greenfie.N, G. D. Fasman, *Biochemistry* **1969**, *8*, 4108; b)J. P. Hennessey, Jr., W. C. Johnson, Jr., *Biochemistry* **1981**, *20*, 1085; c)S. W. Provencher, J. Glockner, *Biochemistry* **1981**, *20*, 33; d)R. W. Woody, *Methods Enzymol* **1995**, *246*, 34.
- [25] M. E. Brandt, L. E. Vickery, *Journal of Biological Chemistry* **1997**, *272*, 4843.
- [26] T. R. Geistlinger, R. K. Guy, *Journal of the American Chemical Society* **2001**, *123*, 1525.
- [27] a)A. C. Gee, J. A. Katzenellenbogen, *Molecular Endocrinology* **2001**, *15*, 421; b)L. A. Luck, J. L. Barse, A. M. Luck, C. H. Peck, *Biochemical and Biophysical Research Communications* **2000**, *270*, 988; c)N. Greenfield, V. Vijayanathan, T. J. Thomas, M. A. Gallo, T. Thomas, *Biochemistry* **2001**, *40*, 6646.
- [28] D. M. Heery, E. Kalkhoven, S. Hoare, M. G. Parker, *Nature* **1997**, *387*, 733.
- [29] R. van Beuningen, H. van Damme, P. Boender, N. Bastiaensen, A. Chan, T. Kievits, *Clinical Chemistry* **2001**, *47*, 1931.
- [30] S. Folkertsma, P. I. van Noort, A. de Heer, P. Carati, R. Brandt, A. Visser, G. Vriend, J. de Vlieg, *Molecular Endocrinology* **2007**, *21*, 30.

- [31] a)A. C. Pike, A. M. Brzozowski, R. E. Hubbard, T. Bonn, A. G. Thorsell, O. Engstrom, J. Ljunggren, J. A. Gustafsson, M. Carlquist, *EMBO J* **1999**, *18*, 4608; b)E. S. Manas, R. J. Unwalla, Z. B. Xu, M. S. Malamas, C. P. Miller, H. A. Harris, C. Hsiao, T. Akopian, W. T. Hum, K. Malakian, S. Wolfrom, A. Bapat, R. A. Bhat, M. L. Stahl, W. S. Somers, J. C. Alvarez, *Journal of the American Chemical Society* **2004**, *126*, 15106.
- [32] a)R. Cudney, S. Patel, K. Weisgraber, Y. Newhouse, A. McPherson, *Acta Crystallogr D Biol Crystallogr* **1994**, *50*, 414; b)J. Jancarik, W. G. Scott, D. L. Milligan, D. E. Koshland, Jr., S. H. Kim, *J Mol Biol* **1991**, *221*, 31.
- [33] W. Kabsch, *Journal of Applied Crystallography* **1993**, *26*, 795.
- [34] A. J. McCoy, R. W. Grosse-Kunstleve, L. C. Storoni, R. J. Read, *Acta Crystallogr D Biol Crystallogr* **2005**, *61*, 458.
- [35] a)A. Warnmark, E. Treuter, J. A. Gustafsson, R. E. Hubbard, A. M. Brzozowski, A. C. Pike, *Journal of Biological Chemistry* **2002**, *277*, 21862; b)A. K. Shiau, D. Barstad, P. M. Loria, L. Cheng, P. J. Kushner, D. A. Agard, G. L. Greene, *Cell* **1998**, *95*, 927; c)R. T. Nolte, G. B. Wisely, S. Westin, J. E. Cobb, M. H. Lambert, R. Kurokawa, M. G. Rosenfeld, T. M. Willson, C. K. Glass, M. V. Milburn, *Nature* **1998**, *395*, 137.
- [36] U. Jonsson, L. Fagerstam, B. Ivarsson, B. Johnsson, R. Karlsson, K. Lundh, S. Lofas, B. Persson, H. Roos, I. Ronnberg, et al., *Biotechniques* **1991**, *11*, 620.
- [37] N. S. Berrow, D. Alderton, S. Sainsbury, J. Nettleship, R. Assenberg, N. Rahman, D. I. Stuart, R. J. Owens, *Nucleic Acids Res* **2007**, *35*, e45.
- [38] L. A. Carpino, G. Y. Han, *Journal of the American Chemical Society* **1970**, *92*, 5748.
- [39] a)J. C. Sheehan, G. P. Hess, *Journal of the American Chemical Society* **1955**, *77*, 1067; b)J. Hachmann, M. Lebl, *Biopolymers* **2006**, *84*, 340.
- [40] a)I. Coin, P. Schmieder, M. Bienert, M. Beyermann, *Biopolymers* **2007**, *88*, 565; b)V. Cavallaro, P. E. Thompson, M. T. W. Hearn, *Journal of Peptide Science* **2001**, *7*, 529.
- [41] K. Mullis, F. Faloon, S. Scharf, R. Saiki, G. Horn, H. Erlich, *Cold Spring Harb Symp Quant Biol* **1986**, *51 Pt 1*, 263.
- [42] J. D. Oliner, K. W. Kinzler, B. Vogelstein, *Nucleic Acids Research* **1993**, *21*, 5192.
- [43] H. Towbin, T. Staehelin, J. Gordon, *Proc Natl Acad Sci U S A* **1979**, *76*, 4350.
- [44] B. W. Matthews, *Journal of Molecular Biology* **1968**, *33*, 491.
- [45] P. Emsley, K. Cowtan, *Acta Crystallogr D Biol Crystallogr* **2004**, *60*, 2126.
- [46] G. N. Murshudov, A. A. Vagin, E. J. Dodson, *Acta Crystallogr D Biol Crystallogr* **1997**, *53*, 240.
- [47] S. Bailey, *Acta Crystallographica Section D-Biological Crystallography* **1994**, *50*, 760.

Chapter 5

Exploring the Role of the Phosphorylated Estrogen Receptor α Ligand Binding Domain via Semi-Synthesis

Abstract: Protein phosphorylation is a versatile posttranslational modification that can regulate nuclear receptor function. The Estrogen Receptor α (ER α) can be phosphorylated on tyrosine 537. For detailed investigation of a possible influence of this phosphorylation on the interaction with other proteins *in vitro*, expressed protein ligation was used to produce the ER α , specific phosphorylated on tyrosine 537. Further, the activity of this semi-synthetic construct with respect to ligand and cofactor binding could be guaranteed.

5.1 Introduction

Nuclear Receptors (NR) undergo a variety of posttranslational modifications (PTMs) in mammalian cells^[1]. Although the transcriptional activity of NRs is mainly governed by ligand binding, PTMs play an important additional part. The influence of these PTMs on the ligand recruitment and the formation of multiple protein complexes on the molecular level is, however unknown. One of the most important and well-studied PTMs in NRs is the phosphorylation^[2]. Although the precise role of receptor phosphorylation is not fully understood, it appears that this covalent change may affect receptor stability, subcellular localization as well as the interaction with other proteins^[1a]. The estrogen receptor α (ER α), as an example, has several phosphorylation sites^[1a] most of them located in the N-terminal part of the receptor, also called activation function 1 (AF-1) that is important for the ligand independent receptor function. In this chapter a tyrosine phosphorylation site in the ligand binding domain (LBD)^[3] of the ER α , and the AF-2, is investigated. The phosphorylation status of this tyrosine 537 plays a regulating role in the hormone binding, protein dimerization, and cofactor recruitment, as deduced from results in cellular assays^[4]. It was furthermore shown that c-Src kinase activity is regulated by binding of its SH2 domain to the phosphorylated tyrosine 537 of the ER α in cellular studies^[5]. Additionally, the tyrosine phosphorylation could be of interest for the regulation of cofactor recruitment, as it is located in the N-terminus of helix 12 in close proximity of the cofactor binding site, especially, as helix 12 is known to take a modulating position for cofactor binding (Figure 1).

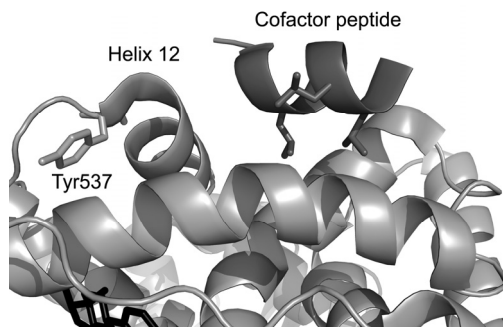


Figure 1: Zoom in on the co-crystal structure of the ER α LBD with the peptide PERM-1 and estradiol (pdb code: 1PCG)^[6]. Tyrosine 537 is located in the N-terminus of helix 12 in close proximity to the cofactor interacting surface.

As the tyrosine 537 phosphorylated ER α construct had yet not been accessible for biochemical studies, detailed molecular information was lacking. Characterizing this specific nuclear receptor phosphorylation site on the molecular level is an important step in elucidating the exact role(s) receptor phosphorylation plays in function. Although traditional enzymatic methods of *in vitro* protein phosphorylation via the c-Src kinase^[7] could be informative, they resulted in only partial and inhomogeneous phosphorylation states like shown before in chapter 4. Expressed protein ligation (EPL) presents a simple and reliable tool to study receptor phosphorylation *in vitro*.

EPL is a useful extension of the native chemical ligation method to ligate two peptides with a N-terminal cysteine and a C-terminal α -thioester, respectively to a native peptide bond in a spontaneous chemoselective reaction^[8]. The EPL allows the functionalization of recombinantly expressed proteins, for example for the introduction of posttranslational modifications^[9]. The development of special expression systems that use self-cleavable intein domains can be used to easily generate recombinant proteins with C-terminal α -thioester groups^[9].

Due to the successful phosphorylation of the ER β LBD of the corresponding tyrosine 488 via semi-synthesis (Chapter 4) it was decided to apply this method for the semi-synthesis of the Y₅₃₇ phosphorylated ER α LBD. This chapter will discuss the successful development of functional homogeneous and site-specific ER α phosphorylation on tyrosine 537 via EPL and reports studies on the influence of the phosphor-group on protein folding and interactions.

5.2 Protein Semi-Synthesis of the Phosphorylated ER α LBD

The ER α LBD had to be separated into two fragments, prior to the incorporation of a phosphorylation of tyrosine 537. One fragment bearing a C-terminal α -thioester and the other one containing a cysteine in the N-terminus. The optimal position for this separation was determined at the natural occurring cysteine 530 in the C-terminus of helix 11 (Figure 2). This would lead to the synthesis of a peptide of 24 amino acids (aa 530 – 553) bearing a cysteine in the N-terminus. This peptide includes the loop between helix 11 and 12 as well as the complete helix 12 and the phosphorylated tyrosine.

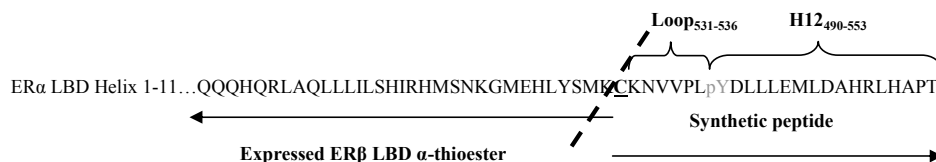


Figure 2: Position of tyrosine 537 in the ER α LBD in close proximity to a cysteine at position 530.

The ER α protein part (aa 302 – 529) was recombinantly expressed as a C-terminal fusion protein with a self-cleavable intein domain^[10] and a chitin binding domain (CBD)^[11] as affinity tag. Further, for biochemical studies a His₈-tag was fused to the N-terminus of the ER α _{302 – 529}. After protein isolation from *E. coli* expression cells, the target protein was purified using chitin affinity chromatography. The intein domain induces an α -thioester linkage which could be cleaved upon addition of an excess of the thiol 2-mercaptoethanesulfonic acid (MESNA)^[12] to generate the truncated ER α _{302 – 529} with a C-terminal α -thioester (His-ER α _{302 – 529} - MESNA) in a concentration of approximately 8 mg L⁻¹. SDS-PAGE and ESI mass spectroscopy analysis (Figure 3) confirmed the presence of a single protein with a mass of 28012.2 Da that corresponds to the calculated mass of His-ER α _{302 – 529} - MESNA (calculated mass: 28005.6 Da).

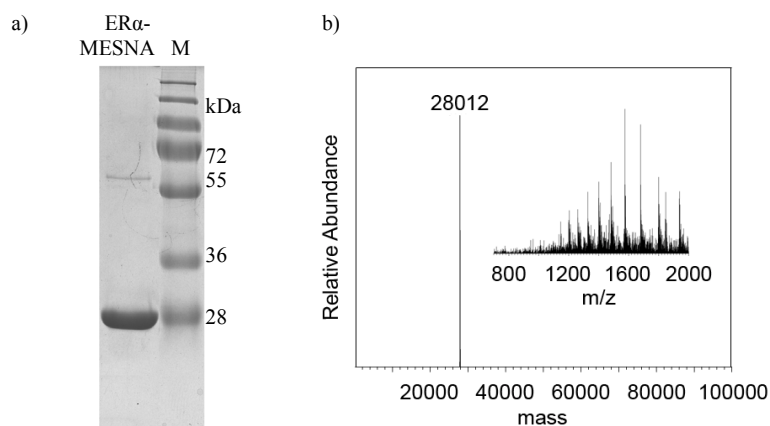


Figure 3: Characterization of the truncated His-ER α LBD₃₀₂₋₅₂₉ α -thioester; (a) SDS-PAGE gel (15 %; Coomassie stained); (b) Deconvoluted mass spectrum of His-ER α LBD₃₀₂₋₅₂₉ - MESNA (calculated mass: 28005.6 Da), inset: m/z spectrum.

The C-terminal part of the ER α LBD including the loop and helix 12 (aa 530-553) was chemically synthesized on solid phase using a peptide synthesizer. Incorporation of the phosphorylation was carried out by using a commercial available tyrosine bearing a phospho-group. For comparison, the non-phosphorylated peptide was synthesized as well. After purification, peptides corresponding to the sequence 530-NH₂-Cys-Lys-Asn-Val-Val-Pro-Leu-Tyr-Asp-Leu-Leu-Leu-Glu-Met-Leu-Asp-Ala-His-Arg-Leu-His-Ala-Pro-Thr-CO₂H-553 and 530-NH₂-Cys-Lys-Asn-Val-Val-Pro-Leu-pTyr-Asp-Leu-Leu-Leu-Glu-Met-Leu-Asp-Ala-His-Arg-Leu-His-Ala-Pro-Thr-CO₂H-553 could be obtained in 20% and 10% of yield,

respectively. The purity (>95% at 220 nm) and molecular weights of the peptides were confirmed by HPLC with ESI-MS detection (Figure 4).

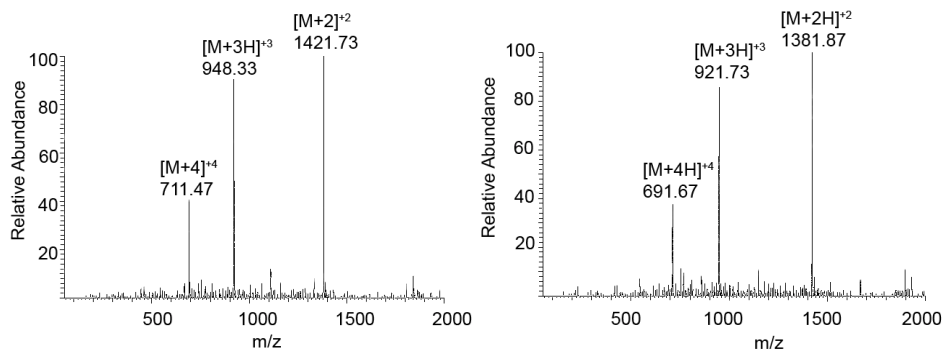


Figure 4: ESI-MS spectrum of $\text{NH}_2\text{-CKNVVPLXDLLLEMLDAHRLHAPT-CO}_2\text{H}$ peptides after preparative HPLC purification; a) Mass spectra with $x = \text{pTyr}$ (calcd. mass 2840.25 Da, detected mass 2841.46 Da); b) Mass spectra with $x = \text{Tyr}$ (calcd. mass 2762.29 Da, detected mass 2761.74).

The native chemical ligation reaction of the His-ER $\alpha_{302-529}$ α -thioester and an excess of pure helix 12 peptide bearing an N-terminal cysteine and a phosphorylated tyrosine and a unmodified tyrosine, respectively was performed in the presence of the thiol MESNA. As the ligation of His-ER $\alpha_{302-529}$ with the phosphorylated peptide increases the molecular weight of the protein by almost 3 kDa, the ligation reaction could be monitored by SDS-PAGE (Figure 5). Typically, the ligation reaction went to completion, without any indication of remaining non-reacted His-ER $\alpha_{302-529}$, when the protein α -thioester was used directly after purification from the chitin resin.

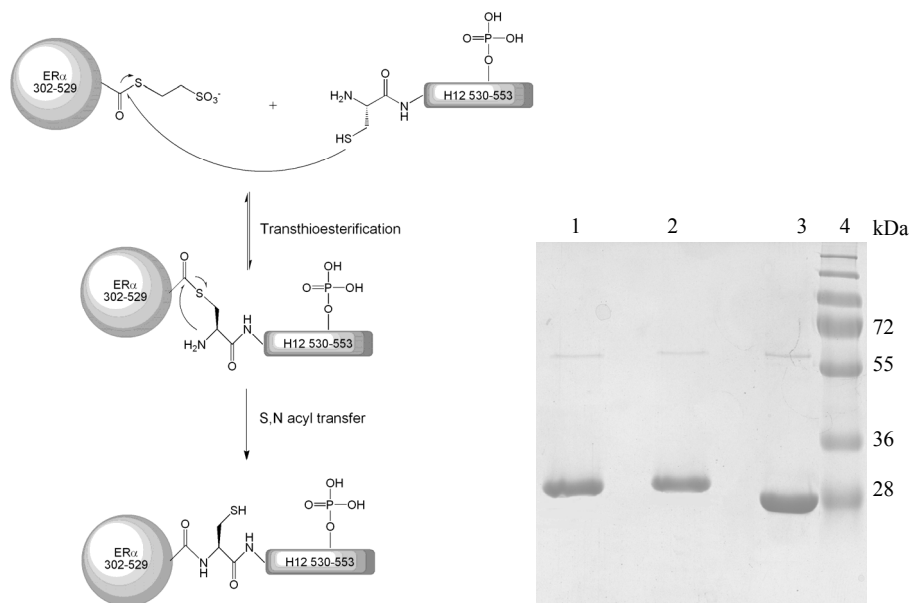


Figure 5: Schematic representation of the native chemical ligation reaction between the truncated His-ER α LBD α -thioester and the helix 12 peptide with an N-terminal cysteine (left). SDS-PAGE gel (15 %; Coomassie stained). Lane 1: phosphorylated His-ER α LBD (calcd. mass: 30722.85 Da), lane 2: non-phosphorylated His-ER α LBD (calcd. mass: 30644.89 Da), lane 3: truncated His-ER α LBD α -thioester (calcd. mass: 28005.6 Da), lane 4: molecular weight marker (right).

The excess of non-bound peptide could be easily removed by immobilized metal ion affinity chromatography (IMAC) using the His-tag in the N-terminus of the fusion protein. In the end, the semi-synthetic ER α LBD could be purified that was homogeneously phosphorylated on tyrosine 537. Similarly, the semi-synthetic ER α LBD with a non-phosphorylated tyrosine could be generated. The non-phosphorylated ER α LBD₃₀₂₋₅₅₃, including helix 12, was also recombinantly expressed in *E. coli* and purified via IMAC using a His₆-tag in its N-terminus. This protein allows for a proper comparison of the functionality of the semi-synthetic proteins with the expressed protein. Pure fractions were collected, combined and analyzed via SDS-PAGE (Figure 6). The impurity with a size of approximately 60 kDa is most probably the bacterial heat shock protein GroEL that strongly binds the nuclear receptor and could not be removed during purification.

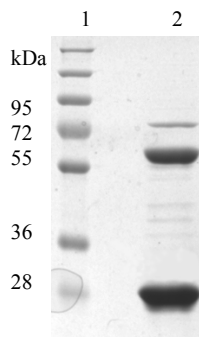


Figure 6: SDS-PAGE gel (15%, Coomassie stained) of heterologous expressed and purified His-ER α LBD (aa 302 – 553). Lane 1: Molecular weight marker, lane 2: Expressed ER α LBD (aa 302 – 553) with a calculated mass of 31059.4 Da.

5.3 Structural Analysis using Circular Dichroism Spectroscopy

The folding of the different ER α LBD protein constructs was evaluated with circular dichroism spectroscopy (CD)^[13]. CD-spectra were monitored from 190 – 250 nm in phosphate buffer at 10 °C. Figure 7 shows the spectra of the non-phosphorylated proteins, generated via semi-synthetic and via expression, and of the phosphorylated ER α LBD in the presence and in the absence of the natural agonist estradiol (E₂). All measurements were performed in three separated experiments to confirm the integrity of the results.

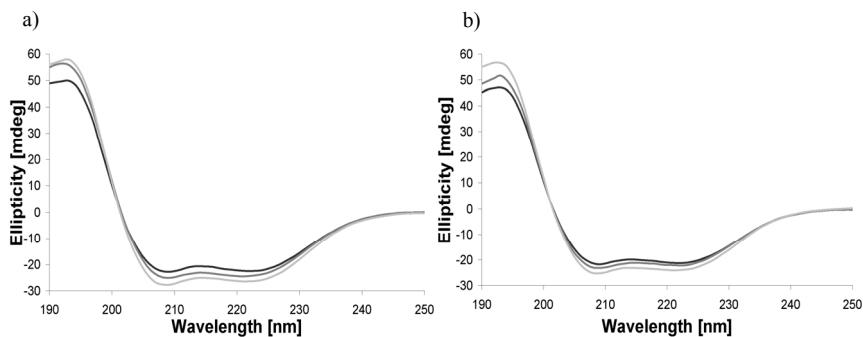


Figure 7: CD spectra of recombinant and semi-synthetic ER α LBDs without ligand (a) and in the presence of estradiol (b). The recombinant ER α LBD is presented in black, the semi-synthetic ER α LBD is presented in dark grey and the phosphorylated ER α LBD with pY₅₃₇ is presented in light grey.

The CD spectra presented in all cases the typical shape of an α -helical secondary structure with a positive CD-signal at a wavelength of around 192 nm and two absorption minima at 208 and 222 nm. Since the ER α LBD consists of twelve α -helices it was expected

that in the correctly folded state the α -helix structural element would dominate the overall CD effect. The CD results thus indicate that the semi-synthetic approach to generate the ER α LBD allows the generation of correctly folded proteins. For better comparison of the three proteins a quantitative estimation of the helical character can be made using the ratio $-\theta_{192\text{nm}} / \theta_{222\text{nm}}$, where a value of 2.63 indicates an ideal helix (Table 1). These data indicate a similar helicity for the reference proteins as well as for the phosphorylated ER α LBD with a value of approximately 2.3. Further, binding of estradiol to the LBDs does not result in significant change in the CD spectra. CD spectroscopy reports only on the overall averaged structural elements, but for example does not determine the specific location of the ER LBD H12. Interestingly, there is no detectable difference between the non-phosphorylated and the phosphorylated ER α LBD. These results are in contrast to the observed decrease in helicity for the phosphorylated ER β LBD (Chapter 4). This implies that phosphorylation of the tyrosine 537 in the ER α LBD has an effect that differs from that function of the phosphorylated tyrosine 488 in the ER β LBD.

Table 1: Overview of the helicity of the recombinant and semi-synthetic proteins determined by the $-\theta_{192\text{ nm}}/\theta_{222\text{ nm}}$ ratios.

Protein/ Additive	Recombinant ER α LBD	Semi- synthetic ER α LBD	Semi-synthetic ER α LBD with pY ₅₃₇
no ligand	2.22	2.32	2.30
E ₂	2.21	2.33	2.36

5.4. Binding Studies using a Cofactor Recruitment On-Chip Assay

Activation of the Estrogen Receptor (ER) typically occurs after binding of an agonist to the ligand binding pocket within the Ligand Binding Domain (LBD) of the receptor. The consequential conformational change, primarily of helix 12, creates a hydrophobic groove on the surface of the ER that can be recruited by transcriptional coactivators via a signature LXXXL motif (where L is a leucine and x is any other amino acid)^[14]. Binding of an antagonist instead induces the interaction with corepressor proteins bearing an LXXXIXXXL motif. In order to investigate the activity of the semi-synthetic proteins (non-phosphorylated and phosphorylated) a cofactor library containing coactivator and corepressor peptides, immobilized on a membrane chip, was screened to detect binding events with the ER α LBD constructs in presence and absence of estradiol (E₂) and the antagonist *trans*-hydroxy-tamoxifen (HT), respectively. To visualize the interactions of the ER α LBD with different

coactivator peptides *in vitro*, we used a fluorescent α -His antibody against the N-terminal His-tag of the recombinant ER α LBD fusion proteins. Binding experiments were performed by incubating the membrane chips with a mixture of antibody, receptor, and ligand, when appropriate (Figure 8). Each spot on the chip represents an individual cofactor peptide immobilized on a porous three-dimensional metal-oxide carrier^[15]. Subsequently, the intensity of ER α binding could be detected via the read out of the fluorescence of the attached antibody. A reference membrane chip was incubated with buffer containing only antibody and ligand to subtract unspecific binding of the antibody and possible buffer effects. Each experiment was performed in triplicate to verify the measured data. The obtained peptide binding profiles are reported in Figure 9.

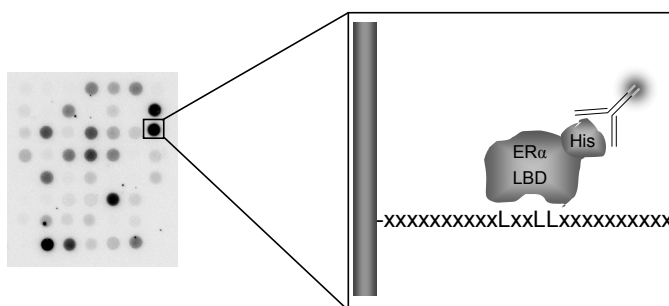


Figure 8: Estrogen receptor-coregulator-interaction profiling. A set of 53 coregulator peptides immobilized on a porous metal-oxide carrier was incubated with the non-phosphorylated and the phosphorylated ER α LBDs, followed by visualizing the binding via a fluorescent α -His antibody. Each spot represents distinct immobilized coregulators. The intensity of each spot resembles fluorescence intensity and thus the amount of bound protein.

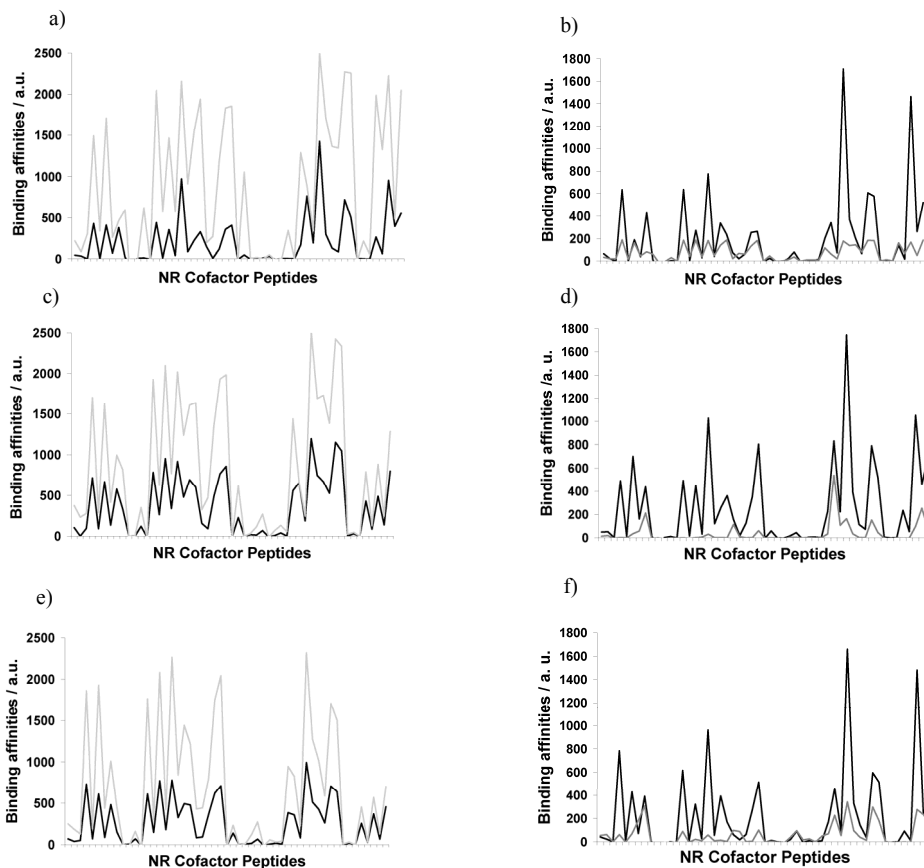


Figure 9: Peptide binding profiles of 53 cofactor peptide sequences immobilized on a membrane against recombinant (a and b), semi-synthetic (c and d) and semi-synthetic ER α LBD with pY₅₃₇ (e and f), in the absence or in the presence of E₂ (left) and HT (right). Protein without ligand is shown in dark grey and protein with ligand (E₂ or HT) is shown in light grey. All protein constructs featured similar binding profiles. Small differences in the dark grey profiles on both sites are due to two independent experiments.

The peptide binding profiles of all three ER α constructs are relatively similar. Typically, peptides that bind with a high affinity to the non-phosphorylated constructs, also bind strongly to the phosphorylated protein. Nevertheless, for selected peptides a slight differentiated binding occurs, as can be deduced from the relative affinities amongst the peptides. Further, the presence of estradiol and *trans*-4-hydroxytamoxifen, respectively had the same influence on the ER α LBDs with respect to cofactor binding. Estradiol increased the basal affinity to coactivator peptides, while incubation with *trans*-hydroxytamoxifen triggered an almost complete loss of the basal binding affinity for most peptides.

The similar peptide binding profiles of the two reference proteins and the phosphorylated LBD is in contrast to the differentiated binding events observed for the phosphorylated ER β LBD in the absence of estradiol (Chapter 4). Apparently, the influence of the tyrosine phosphorylation of the ER LBD is significantly less pronounced in the case of ER α .

In the membrane assay, the peptides are immobilized on a three-dimensional surface that leads to high local concentrations of peptide and possibly to concomitant multivalency effects, that could trigger artificial binding event of the nuclear receptor. Potentially, small effects due to the phosphorylation could be hidden by this multivalency effects.

5.5 Binding Studies using a Cofactor Recruitment FRET Assay

The effect of tyrosine 537 phosphorylation on peptide binding was further studied in solution using a fluorescence resonance energy transfer (FRET) assay. The *in vitro* FRET assay that was applied is based on a technology reported by Folkertsma *et al.* [16] and allows the detection of potential binding events between the non-phosphorylated and phosphorylated ER α LBD, respectively with a selected set of 31 different coactivator peptides in solution. The assay consists of His-tagged ER α LBD, a biotinylated coactivator peptide, an europium-labeled anti-His antibody (Eu), and streptavidin-conjugated allophycocyanin (APC). Upon agonist binding, the streptavidin - biotinylated coactivator peptide complex is recruited by the His-tagged ER α LBD to form a complex resulting in fluorescence resonance energy transfer (FRET) between the europium and allophycocyanin (Figure 10).

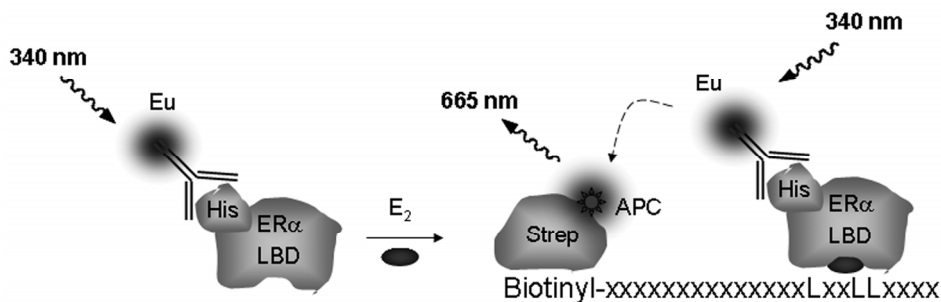


Figure 10: Illustration of the *in vitro* cofactor recruitment assay. Binding of an agonist to the His-tagged ER α LBD will recruit a biotinylated coactivator peptide. Eu-labeled α -His antibody and streptavidin-conjugated APC will assemble into the complex resulting in FRET.

In an initial experiment, the affinity of two different non-phosphorylated ER α LBDs, the semi-synthetic and the expressed constructs, for the coactivator peptide from SCR2 box 1 was

investigated. Increasing concentrations of the peptide were titrated to the protein constructs in the presence of an excess of estradiol (Figure 11). The EC_{50} values of the expressed ER α LBD (EC_{50} : 0.19 μ M) and the semi-synthetic control protein (EC_{50} : 0.17 μ M) are the same. This indicates that the folding and activity of the protein obtained via expression and via semi-synthesis is comparable and proves that expressed protein ligation is a suitable tool to generate semi-synthetic proteins with a native fold. The binding activity of the recombinant, the semi-synthetic and the semi-synthetic ER α LBD with pY₅₃₇ for a specific cofactor peptide (SHP 1) with a fixed concentration was compared to each other depending on increasing concentrations of agonist and antagonist (Figure 12).

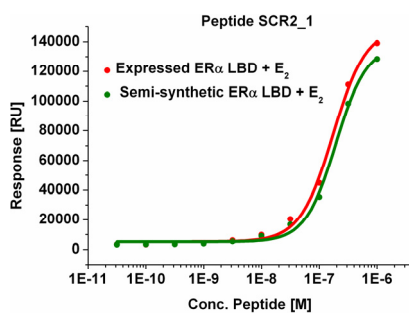


Figure 11: Cofactor recruitment assay using a cofactor peptide from SCR2 box 1. Shown is the binding profile for the recombinant ER α LBD (red) and the semi-synthetic reference protein (green) in the presence of estradiol (E₂). Both proteins feature a very similar affinity to the peptide.

Subsequently, the binding affinity of all three protein constructs to another cofactor peptide (SHP 1) was studied as constant peptide concentration, but with increasing concentrations of either agonist or antagonist (Figure 12). Estradiol (E₂) increased the ability of the cofactor peptide to bind to the ER α LBDs, while the antagonist *trans*-hydroxy-tamoxifen abolished the binding event, as expected. The phosphorylated protein featured a differentiated binding pattern compared to the controls in the presence of E₂, leading to a lower efficiency of the FRET effect. As the two non-phosphorylated proteins exhibited such a high agreement in cofactor affinity it was decided to use only the semi-synthetic ER α as control protein for subsequent investigations of the influence of Y₅₃₇ phosphorylation on LXXLL peptide affinity. Since HT featured no significant change in the binding profile, the differentiation in LXXLL peptide binding affinity was only evaluated in absence of ligand and with the agonist E₂ (Figure 13).

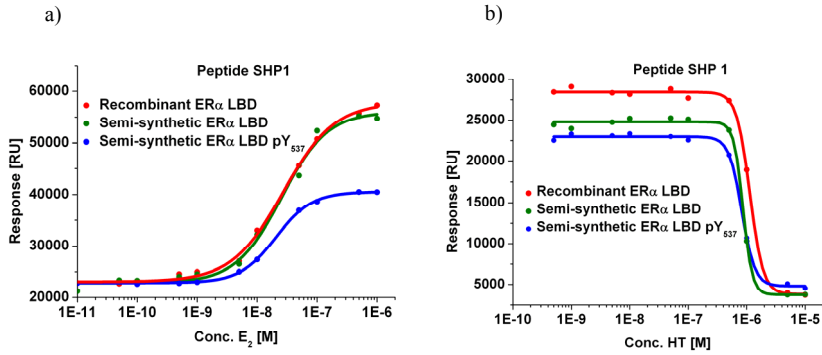


Figure 12: Cofactor peptide recruitment pattern depending on estradiol (E₂; a) and *trans*-hydroxytamoxifen (HT; b), respectively. The two differently generated non-phosphorylated proteins (red and green) featured similar binding activity, while the phosphorylated ER α LBD (blue) showed decreased cofactor affinity with estradiol.

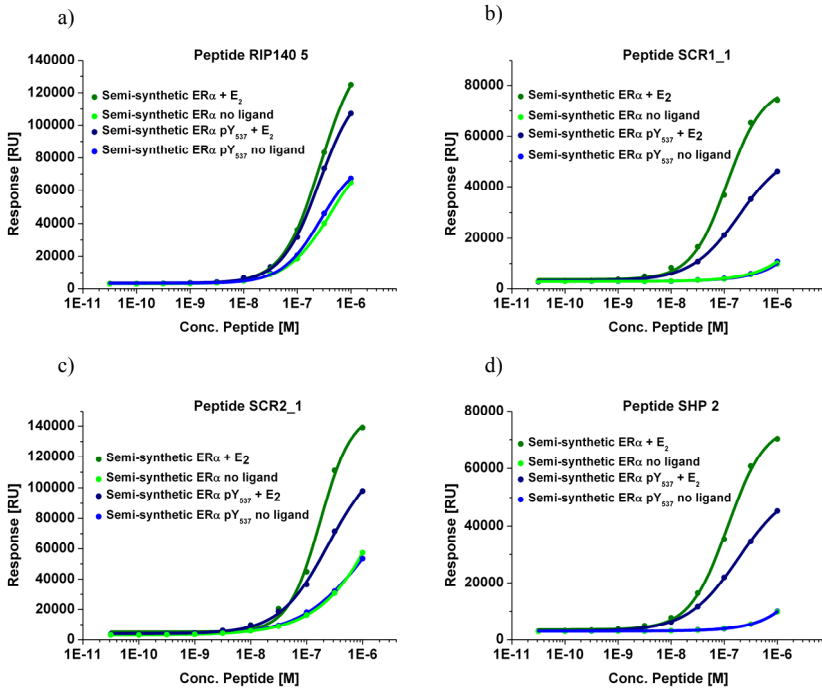


Figure 13: Exemplary binding patterns of a selected set of cofactor peptides for the non-phosphorylated (green) and phosphorylated (blue) ER α LBD in the absence of ligand (pale) and in the presence of estradiol (E₂; dark). Most peptides behaved like the peptide containing the RIP140 box 5 motif (a), where the affinity of the non-phosphorylated and the phosphorylated protein is quite similar in the presence of estradiol and in the absence of ligand. However, for a few peptides (b-d), the phosphorylated ER α LBD features a lower affinity in the presence of estradiol in comparison to the non-phosphorylated reference protein.

The peptide binding results summarized in Figure 13 make clear that a number of peptides differ significantly in their ability to bind the ER α LBDs. In the experiments with estradiol, a saturating concentration was used. The EC₅₀ values thus reflect the affinities of cofactor peptides for the ER α LBDs in their most agonistic conformation (Table 2). The affinity of most cofactor peptides for the non-phosphorylated ER α LBD was similar as for the phosphorylated ER α LBD. However, a limited set of peptides bound the phosphorylated ER α LBD with a lower affinity in the presence of estradiol. For these peptides the basal affinity for the receptor without ligand, however, was independent on phosphorylation state. These efficacy differences in the presence of estradiol may reflect conformational differences in the ER α LBD-coactivator peptide complexes upon phosphorylation, or a specific (repulsive) interaction between the phosphate group and the peptide. Tyrosine 537 phosphorylation thus appears to silence the activating effect of estradiol to recruit specific cofactors. The phosphorylation thus plays the role of a modulator of cofactor recruitment.

Table 2: EC₅₀ values in molar [M], as far as they could be determined, of the different cofactor peptides for the semi-synthetic control ER α LBD and for the phosphorylated ER α LBD in absence of ligand and in the presence of E₂.

Peptide	Non-phosphorylated ER α LBD [M]		Phosphorylated ER α LBD [M]	
	no ligand	E ₂	no ligand	E ₂
SCR2_1	--	1.73E-07	--	2.45E-07
SCR1_2	--	2.35E-07	2.79E-07	3.48E-07
SCR2_3	2.76E-07	2.16E-07	2.79E-07	2.12E-07
SHP_1	--	--	--	--
SHP_2	--	1.15E-07	--	1.86E-07
D22	--	2.61E-07	3.02E-07	3.32E-07
D47	--	--	--	--
DAX3	--	--	3.18E-07	2.93E-07
RIP140_5	--	--	--	--
RIP140_6	--	2.75E-07	2.68E-07	2.62E-07
RIP140_8	--	--	--	--
RIP140_9	--	2.75E-07	--	2.45E-07
CBP_1	--	2.56E-07	--	--
PGC_1	--	1.80E-07	--	1.99E-07
RIP140_3	--	--	--	1.49E-07
HRC α A2	--	1.79E-07	--	--
TRAP220_1	--	--	--	--
TRAP220_2	--	2.95E-07	--	2.77E-07
SRC3_1	--	3.43E-07	--	2.63E-07
SRC3_2	--	--	2.78E-07	--
SRC1a_4	--	1.90E-07	--	3.36E-07
SRC1_3	--	--	--	--
RIP140_1	--	--	3.09E-07	3.37E-07
RIP140_7	--	--	3.03E-07	3.44E-07
D30	--	--	--	3.03E-07
LEEG	--	--	--	--
PERC-1	--	--	--	--
PERC-2	--	--	--	--
EA2	--	--	--	--
ARAF1_4	--	--	--	--
SRC1_1	--	1.15E-07	--	1.91E-07
buffer	--	--	--	--

5.6 Co-Crystallization Studies of the ER α LBD with Estradiol and Cofactor Peptide

X-ray crystallography is an ideal method to get full molecular insight in the effects of tyrosine phosphorylation. Similar to the crystallization of the phosphorylated ER β LBD, as described in chapter 4, the crystallization for the different ER α LBD constructs was therefore envisioned. Experiments were undertaken to solve the x-ray crystal structure of the ER α LBD in complex with estradiol and a cofactor peptide. In order to ensure the ability of a

crystallization of labor-intensive phosphorylated protein, the co-crystallization was first evaluated with the non-phosphorylated expressed ER α LBD. The structure of this protein has previously been solved in complex with various ligands and cofactors^[17].

A peptide was synthesized corresponding to the second interaction motif of the steroid receptor coactivator 1 (SRC1 Box2) with the sequence NH₂-LTERHKILHRLLQEGSPSD-CO₂H. Further, the ER α LBD had to be generated without a final purification tag, in order to avoid a possible negative influence thereof. For this reason, a plasmid expressing the ER α LBD_{289–554} was generated and the protein was expressed in *E. coli*. As a huge part of the ER α LBD existed as inclusion bodies, the protein extract was denaturated via treatment with urea. Refolding, to obtain the biologically active protein was performed by applying the ER α LBD to an estradiol affinity column and subsequently reduction of the concentration of the denaturant. Further, carboxymethylation of all surface cysteines was performed over night at 4 °C using iodoacetic acid, while the protein was bound to the column. Following, the ER α LBD could be eluted with increasing concentrations of estradiol and concentrated to approximately 12 mg/ml. To remove smaller aggregations size exclusion chromatography (SEC) were performed and the purity of the ER α LBD was subsequently controlled by SDS-PAGE (Figure 14).

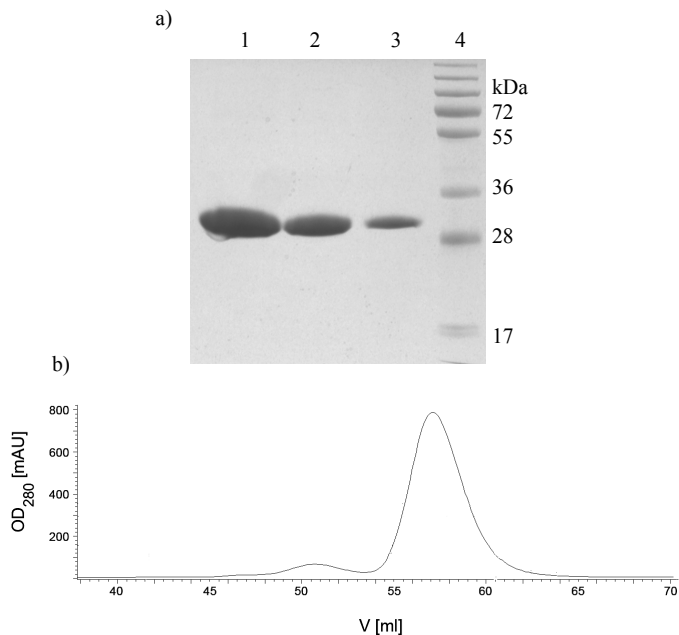


Figure 14: Analysis of the purification of the recombinant ER α LBD; a) SDS PAGE gel (15%, Coomassie stained) of the pure ER α LBD. Lane 1: ER α LBD after elution from the estradiol affinity column using 100 μ M E₂, lane 2: using 150 μ M E₂ and lane 3: using 200 μ M E₂. Lane 4: Molecular weight marker; b) Elution profile of the expressed ER α LBD (calcd. mass: 29336.5 Da) after SEC-analysis using FPLC.

The pure ER α LBD/E₂ complex was combined with the SRC1 box 2 peptide, dialyzed against the protein buffer, using a 1.5 or 3-fold molar excess. A screening of suitable crystallization conditions was carried out for the ER α LBD/E₂/peptide complex using variable concentrations of protein and peptide. Crystallization was performed in the sitting drop fashion using equal volumes of the ER α LBD/E₂/peptide complex and reservoir solution were equilibrated against reservoir solution at 20 °C and at 4°C. After two days, promising thin rod-shaped crystals appeared in one condition (0.1 M Bicine, pH 9, 20% PEG 6000). These crystals were crooked and did not grow beyond 90 μm x 17 μm x 5 μm after 21 days and could not be reproduced for fine screening applications (Figure 15). However, after soaking one of these crystals in reservoir solution complemented with glycerol to 20%, the crystal diffracted to about 4 Å. The achieved resolution was not sufficient enough to determine the structure of the protein. A further problem was the anisotropic behavior of the thin crystal pads in the X-ray.



Figure 15: Co-Crystals of the recombinant ER α LBD complexed with E₂ and cofactor peptide under conditions of 0.1 M Bicine pH 9 and 20% PEG6000 (v/v) at 4 °C.

The ER α LBD is known to aggregate in the absence of reducing agents, which could complicate handling and inhibit the formation of highly ordered crystals. Reducing agents (DTT, TCEP, β -mercaptoethanol) were therefore added to the protein buffer, however this did not lead to better protein crystals. In order to circumvent the cysteine carboxymethylation with iodoacetic acid and possible side-effects, all surface cysteines were replaced by serines via mutagenesis. Based on studies of scientists at Bayer-Schering-Pharma (personal communication with Vera Pütter), cysteine 381 and cysteine 417 in the ER α LBD were chosen for mutagenesis. However, crystallization of the resulting mutant Δ ER α LBD₂₉₈₋₅₅₄ C381S/C417S caused neither improvement of the crystal quality.

A different purification strategy, disclaiming the use of the estradiol affinity column and also avoiding the refolding of the protein from the inclusion bodies using urea, was applied to purify the cysteine to serine mutant ER α LBD. The protein was provided with an

N-terminal His₆-tag and after expression, first the $\Delta ER\alpha$ LBD₃₀₂₋₅₅₄ C381S/C417S was purified from the soluble *E. coli* protein extract via IMAC, followed by the cleavage of the His₆-tag and gel filtration. Subsequent anion exchange chromatography (AEX) separated the protein into two peaks. On a SDS-PAGE gel the two peaks showed discrete bands (Figure 16). As both bands featured the expected mass of 29.3 kDa both fractions were incubated with estradiol and cofactor peptide and subsequently screened for suitable crystallization conditions. No crystals were obtained for protein eluting in the second AEX peak. Crystallization of protein from the first AEX resulted in promising, but somewhat crooked small crystals that could be reproduced (Figure 17). These crystals appeared after 4-6 days and grew to dimensions of 120 μm x 20 μm x 6 μm within 35 days. Although these crystals were still too small to be measured, streak-seeding was performed in order to improve the growth of these crystals. The crystals were transferred to the center of a pre-equilibrated drop using a whisker. The recipient drop contained a solution of protein and mother liquor with 10% less precipitant. Streak-seeding could increase the size of the crystals up to 300 μm x 50 μm x 15 μm , but also increased the adhesions. Although the crystals were still fairly crooked, they diffracted beyond 4 \AA . However, the quality of the crystals was still not sufficient enough to analyze then via x-ray. Future studies need to concentrate on other crystallization conditions with special additives like detergents, denaturants, alcohols, salts and pH.

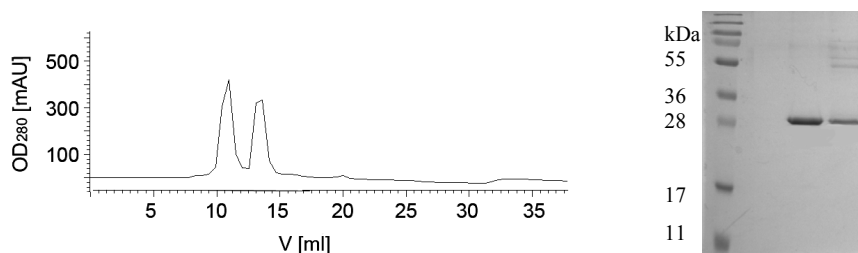


Figure 16: Anion exchange chromatography separated the $\Delta ER\alpha$ LBD C381S/C417S into two peaks. On the SDS-PAGE gel both peaks featured discrete bands of the same size. Calculated mass of the $\Delta ER\alpha$ LBD: 29304.4 Da.

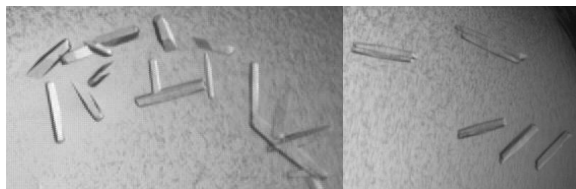


Figure 17: Co-Crystals of the expressed ER α LBD complexed with E₂ and cofactor peptide, grown under conditions of 0.2 M Potassium Fluoride and 20% PEG3350 (v/v) at 4 °C (left) and 0.2 M Sodium Formate and 20% PEG3350 (right).

5.7 Binding Studies using Surface Plasmon Resonance (SPR)

The phosphorylation of tyrosine 537 has been shown to lead to interactions of the ER α with the SH2 domain of the c-Src kinase in cellular assays^[5]. In these studies it was also shown that this protein-protein interaction is regulated by agonist binding to the receptor ligand binding pocket. In order to investigate a potential interaction of the SH2 domain with the phosphorylated tyrosine 537 *in vitro*, a surface plasmon resonance (SPR)^[18] assay was performed. Different concentrations of the non-phosphorylated and the phosphorylated ER α LBD were flown over the immobilized c-Src SH2 domain, followed by dissociation in the same buffer without protein. Unfortunately, a binding of the phosphorylated and the non-phosphorylated ER α LBD to the surface was observed (Scheme 18). This indicates that the observed signal increase occurs not due to specific binding of the ER α LBD to the SH2 domain via the phosphorylated tyrosine at position 537. Apparently the current immobilized protein surface leads to unspecific absorption. This might be overcome with the usage of other immobilization techniques or different protein constructs. Also the study of this protein-protein interaction via isothermal titration calorimetry (ITC) and fluorescence polarization (FP) using a fluorescent labeled SH2 domain, could provide detailed insight on the interaction of c-SRC kinase with the ER α LBD.

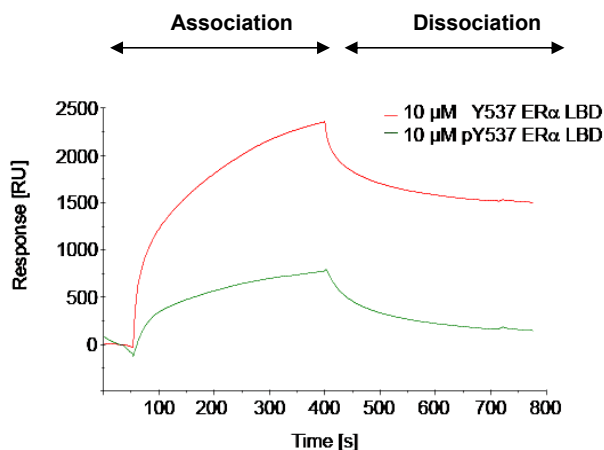


Figure 18: Sensogram of the interaction of the immobilized c-Src SH2 protein construct with 10 μM of the phosphorylated ER α LBD (green) and with 10 μM non-phosphorylated ER α LBD (red) using SPR.

5.8 Conclusions

Expressed protein ligation could be successfully established as a technique to generate an ER α LBD that is homogeneously and selectively phosphorylated on tyrosine 537. It could be demonstrated that this semi-synthetic protein is correctly folded and active with respect to ligand and cofactor binding. No differences in behavior were observed between the non-phosphorylated expressed and non-phosphorylated semi-synthetic ER α LBDs, confirming the valid entry into these proteins via semi-synthesis.

Phosphorylation on tyrosine 537 is able to decrease the binding efficiency of distinct coactivator peptides to the agonist-bound receptor, while having a minimal effect on cofactor affinity of the antagonist-bound and apo-ER α LBD, respectively. This observation strongly suggests that tyrosine phosphorylation of the ER α LBD represents an important control site that is involved in regulating cofactor binding under certain cellular conditions. This ligand-dependency mirrors the previously reported effect of tyrosine phosphorylation of the ER β LBD on cofactor binding (Chapter 4). In contrast to the phosphorylation-mediated enhancement of cofactor binding activity in the absence of ligand for the ER β LBD, however, phosphorylation of the ER α LBD did not influence ligand independent cofactor binding, but rather modulated the affinity of specific peptides for the receptor surface in the presence of estradiol. This differentiated effect of tyrosine phosphorylation between the two ER receptor subtypes reinforces the idea that phosphorylation provides the two ER subtypes with distinct cofactor regulatory functions.

An additional role of the tyrosine phosphorylation of the ER α LBD lies in the control over other protein-protein interactions via the phosphorylated tyrosine. One of these supposed ER interacting proteins is the SH2 domain of the c-Src kinase. In this current study, interaction of SH2 with the ER α LBD via tyrosine phosphorylation could not be evinced by yet, but the generated proteins provide a good platform to study these effects in detail with a variety of characterization techniques. Similarly, optimization of the protein crystallization efforts could lead to the solving of the crystal structure of the phosphorylated ER α LBD, possibly in complex with regulatory proteins, and will give deeper insights in how phosphorylation influences ER α function.

5.9 Experimental Section

General. Unless stated otherwise, all reagents and chemicals were obtained from commercial sources and used without further purification. All chemical reagents were purchased from Novabiochem, Aldrich-Sigma, Fluka and Acros. All biological reagents were purchased from Invitrogen, Fermentas, New England Biolabs, Qiagen, Novagen, Stratagen, Amersham Bioscience, Serva, Sigma-Aldrich, Thermo Scientific and Pharmacia. LC-ESI-MS was carried out by using an Agilent 1100 series binary pump together with a reversed phase HPLC C18 column (Macherey-Nagel) and a Finnigan Thermoquest LCQ. If not otherwise stated, the following gradient program was used for analytical LC-MS: flow: 1 mL/min, solvent A: 0.1% HCO₂H in H₂O, solvent B: 0.1% HCO₂H in CH₃CN, A/B: 90/10 (0-1 min) to 0/100 (over 10 min). Purification of products by RPHPLC was performed in an Agilent 1100 Series Purification Platform using a NUCLEODUR® C18 Gravity preparative column from Macherey-Nagel (21 x 250 mm) and flow rate of 25 ml/min. The products were eluted by using different solvent gradients of solvents A and B (solvent A = 0.1% TFA/H₂O; solvent B = 0.1% TFA/CH₃CN). UV signal at 210 nm was used for detection.

Primers used for all the cloning procedures were supplied by MWG Biotech (Ebersberg, Germany) and all restrictions enzymes were purchased from New England Biolabs (Ipswich, MA, USA) or Fermentas (Mannheim, Germany).

Peptide synthesis. Fmoc-protected amino acids^[19] were purchased from MultiSyntech and Novabiochem in their appropriately protected form. Solid phase peptide synthesis^[20] was carried out from the C- to the N-terminus with a Syro II automated peptide synthesizer (MultiSynTech GmbH) utilizing Fmoc-Gly-Wang resin with a loading of 0.78 mmol/g (Novabiochem) and Fmoc-Rink amide MBHA resin (Novabiochem) with a loading of 0.7 mmol/g, respectively and standard Fmoc-protected amino acids; the average batch size was 93.6 mmol. The α -amino Fmoc protecting group was removed using 20% piperidine in DMF. Double amino acid coupling mixtures were prepared by dissolving 4 equivalents of the appropriate amino acid in a mixture of HOBt^[21] and addition of a solution of DIC. After completion the cleavage from resin and the final side chain deprotection was achieved with a mixture of TFA/H₂O/EDT/TIS (94:2.5:2.5:1)^[22] after a reaction time of 2 h. Peptides corresponding to the non-phosphorylated NH₂-CKNVVPLYDLLLEMLDAHRLHAPT-CO₂H and the phosphorylated ER α LBD NH₂-CKNVVPLpYDL-LLEMLDAHRLHAPT-CO₂H were obtained in 10% and 20% of yield, respectively and a purity of >95%. Peptides for the Src 1 Box 2 peptide with the sequence NH₂-

LTERHKILHRLQLQEGSPSD-CO₂H was obtained in 28% yield and a purity of >99%. Incorporation of the phosphorylated Y537 was achieved using the Fmoc-Tyr-(PO(NMe)₂)-OH building block, which was hydrolyzed to the corresponding phosphate by dissolving the peptide cleaved from the resin in a mixture of TFA/H₂O (1:1) overnight at 4 °C. The peptides were precipitated and washed three times by ice-cold diethylether followed by purification with preparative HPLC on a C18 column using a H₂O-ACN gradient and lyophilization. The purity (>95%) and molecular weights of the final compounds was determined using analytical HPLC.

Polymerase Chain Reaction. PCR^[23] was performed using a T3000 thermocycler (Biometra) or a Mastercycler epgradient (Eppendorf). The general PCR preparation was consisted of a total volume of 50 µl containing 1 x *Pfu* DNA polymerase reaction buffer (Fermentas), 5-50 ng dsDNA template, 125 ng of each primer (MWG Biotech) and 200 µM dNTP mix (Fermentas) and 2.5 U/µl *Pfu* DNA polymerase (Fermentas). The amplification of DNA resulted from an initial denaturation at 95 °C for 5 min followed by 30-40 cycles of 30 s denaturation at 95 °C, 45 s of hybridization at specific temperatures (5°C lower then melting temperatures of the primers) and DNA synthesis for 2 min at 72 °C. The last step was a terminal DNA synthesis for 5 min at 72 °C. Control and purification of the amplification product was performed by agarose gel electrophoresis (0.8% agarose) in 40 mM Tris/Acetate pH 7.6, 1 mM EDTA, and SYBR Safe DNA gel stain (Invitrogen). Isolation of DNA from the gel was performed using the Qiaquick PCR purification kit or the Qiaquick gel extraction kit (Qiagen). Isolation of plasmid DNA from *E. coli* was performed using the QIAprep Spin Miniprep Kit (Qiagen).

Cloning, expression and purification.

ER α LBD with an N-terminal cleavable His₆-tag: The plasmid pET15-hER α LBD expressing hER α LBD (residues 302-553) with an N-terminal His₆-tag was a kind gift of Arie Visser (Organon, now Merck).

The plasmid was transformed into a chemical competent high-density culture of *E. coli* Rosetta 2 (DE3) host cells (Novagen) that were grown in 4 L TB medium using selection with Ampicillin (100 µg/ml). The cultures were incubated at 37 °C till an OD_{600nm} of ~1.2 and after cooling down to 18 °C protein expression was induced by adding isopropyl β -D-1-thiogalactopyranoside (IPTG) to a final concentration of 100 µM. The cells were grown for additional 18-20 h at 15 °C and if necessary 17- β -Estradiol (E₂, Serva) was added to a final concentration of 10 µM. Subsequently cells were harvested by centrifugation (Beckman Coulter, Avanti J26 XP; 4500 rpm, 20 min, 4 °C) and dissolved in a buffer of 1xPBS, 370 mM NaCl, 40 mM Imidazol, pH 8 and 10 % glycerol (10 ml/g of cells) or stored at -80 °C till further use.

Subsequently, the bacterial suspension was lysed by up to four cycles under a pressure of 80 bar via shear forces using a microfluidizer (Microfluidizer 1109) and centrifuged (Beckman Coulter, Avanti J25; 20.000 rpm, 30 min, 4 °C). The soluble cell lysate was immobilized on an equilibrated Nickel-NTA agarose column (HisTrap HP, 5 ml, Amersham Biosciences) using Fast Protein Liquid chromatography (Äkta FPLC, Amersham Biosciences) at a flow rate of 1 ml/min and washed with lysis buffer to remove non-specific binding. The His₆-ER α LBD fusion protein was finally eluted via an imidazole gradient using elution buffer (1x PBS, 370 mM NaCl, 500 mM Imidazol pH 8 and 10% glycerol). Fractions containing the fusion protein were combined and desalted on a Sephadex G25 PD-10 column (Amersham Biosciences) using desalting buffer (20 mM Tris, 25 mM NaCl, 10 % glycerol and 0.05% β -Octylglycosid). Purity and characterization of the eluted fractions was established by SDS-PAGE using a molecular weight marker (Page Ruler Plus Prestained Protein Ladder, Fermentas) and photometric determination of protein concentration using Nanodrop at a wavelength of 280 nm.

ER α LBD for crystallization: ER $\alpha_{289-554}$ was amplified from pET15-hER α by PCR using the forward primer 5'-GAATTCTCATGATCAAACGCTCTAAGAAGAACAGCC-3' and the reverse primer 5'-TTTTTCTCG-AGTTAGCTAGTGGGCGCATGTAGGCGGT-3' to introduce the restriction site for *RcaI* (*Bsp*HI) and *XhoI*. The amplification product was gel purified, digested with *RcaI* (*Bsp*HI) and *XhoI* and inserted into the *E. coli* expression vector pET16b (Novagen) that has been double digested with the endonucleases *RcaI* (*Bsp*HI) and *XhoI*.

Transformants containing the correct plasmid sequence were used for ER α LBD expression. The ER $\alpha_{289-554}$ was over-expressed from a high-density culture of *E. coli* Rosetta 2 (DE3) host cells (Novagene) and had the following sequence: M[I₂₈₉-S₅₅₄]. Harvested cells were disrupted by passing 5 times through a microfluidizer (Microfluidizer 1109) in a buffer containing 100 mM Tris-HCl pH 8.5, 100 mM KCl, 4 mM DTT, and 1 mM EDTA (10 ml/g of cells). After centrifugation (20.000 rpm, 30 min 4 °C), the pellet was extracted with 4 M urea in the same buffer. The urea extract was applied to a pre-equilibrated 3 ml estradiol-Sepharose column^[24] (PTI Research, Inc.) and first washed with 50 ml 1 M urea in the above buffer and then sequentially washed with the following: 50 ml of 50 mM Tris-HCl, pH 8.5 containing 700 mM KCl, 1 mM DTT and, 1 mM EDTA, followed by 50 ml of 50 mM Tris-HCl pH 8.5, 250 NaSCN, 1 mM EDTA, and 1 mM DTT in 10% dimethylformamide (DMF). Subsequently the column was washed with 50 ml of 10 mM Tris-HCl, pH 8.0. While the ER α LBD was bound to the estradiol-affinity column, carboxymethylation was performed by equilibrating the column with 5 mM iodoacetic acid in 10 mM Tris-HCl pH 8.0 overnight at 4 °C. Elution of the protein was performed with 100-200 μ M estradiol in 10 mM Tris-HCl pH 8.0, followed by desalting into 50 mM ammonium bicarbonate pH 7.5 using a disposable PD10-column (GE Healthcare Life Sciences).

Subsequently partial contaminations were removed by size exclusion chromatography (Sephadex 75, HiLoad 26/60, Pharmacia) using 50 mM ammonium bicarbonate pH 7.5. Pure fractions containing the purified ER $\alpha_{289-554}$ were recombined and concentrated to approximately 12 mg/ml using Amicon ultra centrifuge tubes (Millipore, MWCO 10 kDa) and characterized by SDS-PAGE and photometric determination of protein concentration using Nanodrop at a wavelength of 280 nm.

Δ ER α LBD₂₉₈₋₅₅₄ C381S/C417S: The cysteine to serine mutations were introduced in two steps using the QuickChange Site-directed mutagenesis kit (Stratagene). For the C381S mutation pET16-hER $\alpha_{289-554}$ was used as template for the PCR with the forward primer 5'- CAGGTCCACCTTCTAGAATCTGCCTGGCTAGAGAT-CCTG-3' and the reverse primer 5'- CAGGATCTCTAGCCAGGCAGATTCTAGAAGGTGGACCTG-3'. Digestion of the non-mutated template DNA was achieved by incubating the amplification reaction for 1 h at 37 °C with the endonuclease *DpnI* (10U/ μ l). After transformation of the DNA into XL1-Blue competent cells (Stratagene) the correct open reading frame was confirmed by DNA sequencing (MWG Biotech AG). The C417S mutation was introduced into the Δ pET15-hER $\alpha_{289-554}$ C381S by PCR using the forward primer 5'- GACAGGAACCAGGGAAAATCTGTAGAGGGCATGGTGGAG-3' and the reverse primer 5'- CTCCACCA-TGCCCTCTACAGATTTCCCTGGTTCCTGTG-3' to yield Δ pET15-hER $\alpha_{289-554}$ C381S/C417S. Expression in *E. coli* Rosetta 2 (DE3) cells and subsequent purification was performed like above. Introduction of both mutations into the DNA expressing for the ER α LBD α -thioester with N-terminal His-tag and without tag was performed in the same way.

His- Δ ER α LBD₃₀₂₋₅₅₄ C381S/C417S: The mutated ER α LBD bearing an N-terminal His-tag was established by using the pET15-hER α LBD as a template for the PCR. The same primers like above mentioned were used. After *DpnI* digestion and transformation into *E. coli* XL1 Blue cells (Stratagene) the success of the mutagenesis was confirmed by DNA sequencing (MWG Biotech AG).

Expression of the His- Δ ER α LBD₃₀₂₋₅₅₄ C381S/C417S was performed in Rosetta 2 (DE3) *E. coli* cells that were grown in 4 L TB medium supplemented with 10 μ M E₂ and 10% sucrose at 37 °C till an OD_{600nm} of ~1.2. Expression of the target protein was induced by 100 μ M IPTG and cells were grown for additional 6 h at 25 °C. Cells were harvested by centrifugation (Beckman Coulter, Avanti J26XP; 4500 rpm, 20 min, 4 °C) and dissolved in ice-cold lysis buffer containing 20 mM Tris-HCl pH 8.0, 500 mM NaCl, 10% glycerol, 40 mM imidazol, 0.05% n-Octyl- β -D-glucopyranosid (Roth), 5mM DTT, 10 μ M E₂ and 20 μ M PMSF. Cell disruption was carried out by passing the cell extract four times through a microfluidizer (Microfluidizer 1109). After centrifugation (Beckman Coulter, Avanti J25; 22.000 rpm, 30 min, 4 °C), the supernatant was loaded on a pre-equilibrated Ni²⁺-NTA column (HisTrap HP, 5 ml, Amersham Biosciences) using the Äkta FPLC (Amersham Biosciences) or the Äkta explorer (GE Healthcare) with a flow rate of 1 ml/min. The column was washed with 100 ml of lysis buffer without PMSF at a flow rate of up to 4 ml/min to remove non-specific binding, followed by elution of the ER α LBD in an imidazol gradient from 0 to 500 mM at 1 ml/min. Fractions of interest were pooled and analyzed by SDS-PAGE. Fractions containing the pure fusion protein were combined and the buffer was exchanged to 20 mM Tris-HCl pH 8.0, 300 mM NaCl, 10% glycerol, 0.05% n-Octyl- β -D-glucopyranosid, and 1 mM DTT using PD-10 desalting columns (GE Healthcare Life Sciences). After thrombin cleavage of the His₆-tag, the mutated ER α LBD was concentrated to 10 mg/ml (Amicon Ultra, Millipore, MW 10 kDa) in a buffer of 10 mM Tris-HCl pH 8.0, 10 mM NaCl, 10 μ M E₂ and 50 mM β -Mercaptoethanol. Subsequently the protein was applied to a pre-equilibrated anion exchange chromatography (AEX) column (HiTrap Q FF, 5 ml, Amersham Biosciences) at a flow rate of 1ml/min. A salt gradient from 10 mM to 2 M NaCl over 20 column volumes (CV) was used for elution. Fractions were analyzed on SDS-PAGE and separated into two different pools, which were processed separately. The protein was concentrated to 1-2 ml via Amicon filter (Millipore, MW 10 kDa) before gel filtration (HiLoad 26/60 Sephadex 75). The column was equilibrated with a buffer of 50 mM Tris-maleate pH 8.0, 50 mM NaCl, 10 μ M E₂, and 50 mM β -Mercaptoethanol. Peak fraction with an apparent molecular weight of a dimer were pooled and concentrated by ultrafiltration (Amicon, Millipore, MW 10 kDa) to 12 mg/ml for crystallization trials.

ER α LBD α -thioester: For the construction of the ER α ₃₀₂₋₅₂₉-intein-CBD expressing plasmid, the ER α ₃₀₂₋₅₂₉ fragment was amplified by PCR from pET15-hER α ₃₀₂₋₅₅₃ using the forward primer 5'- TTTTTCATATGAA-GAAGAACAGCCTGGCCTTGCC-3' and the reverse primer 5'-TTTTTGGCTCTTCTGCACTTCATGCTG-TACAGATGCTCCATGCC-3', introducing the restriction sites for the endonucleases *NdeI* and *SapI*. This fragment was subsequently inserted into a double digested pTWIN1 *E. coli* expression vector (New England Biolabs).

The constructed expression plasmid was transformed into *E. coli* BL21 (DE3) cells and protein expression was performed after cell grows and induction with IPTG. After sedimentation and cell lysis with a microfluidizer in a buffer of 20 mM HEPES/NaOH (pH 7,3), 500 mM NaCl, 1 mM EDTA, 10% glycerol, 20 μ M PMSF, 0.1 mM TCEP, 0.1% Triton-X100 and if necessary 100 μ M E₂ the protein extract was centrifuged again. The soluble protein suspension was subsequently applied to a 30 ml pre-equilibrated chitin bead column (New England

Biolabs) at 4°C and washed with 10 column volumes of buffer containing 20 mM HEPES/NaOH (pH 7.3), 1 M NaCl, 1 mM EDTA, 0.1 mM TCEP, 10 % glycerol, 1mM ATP, 2 mM Mg₂Cl and if necessary 10 μ M E₂. To remove ATP, MgCl and the excess of NaCl the column was washed with two column volumes of buffer containing 20 mM HEPES/NaOH (pH 7.3), 500 mM NaCl, 1 mM EDTA, 0.1 mM TCEP, 10 % glycerol and if necessary 10 μ M E₂. On-column cleavage of the intein-tag was induced by equilibrating the chitin beads with two volumes cleavage buffer (20 mM HEPES/NaOH (pH 8.5), 500 mM NaCl, 1 mM EDTA, 10 % glycerol, 200 mM MESNA and if necessary 10 μ M E₂) for 24 h at 4 °C. Elution of the protein was performed by washing the column with buffer cleavage buffer. Elution fractions were collected and pooled after which the cleavage step was repeated to gain more thioester terminated proteins. The proteins were concentrated using Amicon ultra centrifuge tubes (Millipore, MW 10 kDa) and the efficiency of intein cleavage was confirmed by SDS gel electrophoresis and ESI mass spectroscopy.

ER α LBD α -thioester with an N-terminal His₈-tag: The N-terminal His₈-tag was introduced by subcloning an *AfIII/SapI* fragment of pET15-hER α ₃₀₂₋₅₂₉ into the plasmid pTWIN1-His₈ digested with *NcoI/SapI*. Used primers for the PCR were: 5'-GGCCATACATGTCTGGTAAGAAGAACAGCCTGGCCTT-3' and 5'-TTTT-TTGCTCTTCTGCACTTCATGCTGTACAGATGCTCCATGCC-3'. The pTWIN1-His₈ vector was created by cloning a His₈-tag into the vector pTwin1 using the primers 5'-TATGGAAGCGAGCCACCATCACCATCACCATCACCATG-3' and 5'-CATGGCATGGTGATGGTGATGGTGATGGTGGCTCGCTTCCA-3' and ligating the PCR fragment with the *NdeI/NcoI* digested pTwin1 vector. Expression and purification of the ER α LBD α -thioester with N-terminal His₈-tag was performed like described for the ER α LBD α -thioester without His₈-tag via intein thiolysis.

Cleavage of a His₆-tag: If necessary enzymatic cleavage of the His₆-tag was performed using a thrombin protease (1 unit/ μ l; Amersham Biosciences). One cleavage unit was incubated with 100 μ g of fusion protein in desalting buffer at room temperature over night. After cleavage the thrombin was removed from the sample by addition of Benzamidine sepharose beads (GE Healthcare), which bind thrombin. The beads were three times washed in distilled water before usage. The beads (1 μ l per 1 μ l protein) were incubated with the sample for about 30 minutes at 4 °C under slow stirring followed by a 10 minutes centrifugation step at 4500 rpm to remove the sepharose.

SH2/SH3 domain: The SH2/SH3 domain (aa 84-248) of the c-Src kinase was generated as fusion protein with an N-terminal His₆-tag. The template DNA for the PCR was drawn from the human v-Src sarcoma viral oncogene homolog (Ultimate ORF clone IOH 12563, Invitrogen). Used forward primer: 5'-AAGTTCGTGTTTC-AGGGTGGTGGAGTGACCACCTTTGTGG-3' and used reverse primer: 5'-CTGGTCTAGAAAGCTTCAGC-ACACGGTGGTGAGGCG-3'. The PCR product was gel purified and mixed with the linearized pTriEX-2-based^[25] expression vector pOPINF (Dortmund Protein Facility, DPF), allowing a ligation-free *in vivo* cloning strategy^[26]. After transformation of the DNA mix into *E. coli* One Shot OmniMax cells (Invitrogen) the correct open reading frame was confirmed by DNA sequencing (MXG Biotech AG). Expression of the His-SH2/SH3 domain was performed in *E. coli* BL21 (DE3) Codon + RIL expression cells (Stratagene) overnight at 15 °C using 100 μ M IPTG. Upon cell disruption, the soluble protein extract was applied to a Ni²⁺ NTA column (His Trap FF crude 5 ml, Amersham Biosciences) using the Äkta FPLC. The tagged target protein binds to the matrix

and unbound protein is washed out with a buffer of 50 mM Na₂HPO₄, 300 mM NaCl, 1 mM TCEP and 20 mM imidazol pH 8.0. For cleavage of the His-tag, the column is equilibrated for 5 h with cleavage buffer (50 mM Hepes, 150 mM NaCl and 1 mM TCEP pH 8.0) containing 0.7 CV PreScission Protease (GE Healthcare). Purity of the eluted SH2/SH3 domain, cleaved from the His-tag, was confirmed by SDS-PAGE.

Size exclusion chromatography of the ER α LBD. Size exclusion chromatography was used for preparative purification and analytical confirmation of the purity of proteins. The FPLC (Äkta FPLC, Amersham Biosciences or Äkta explorer, GE Healthcare) was performed with filtrated and degassed buffers. For preparative approaches a HiLoad 26/60 Sephadex 75 column was used while analytical experiments were carried out using a HiLoad 16/60 Sephadex 75 column. The columns were equilibrated with two volumes of target buffer at 4 °C. The protein was concentrated to more than 10 mg/ml in a volume of ~1.5 ml and applied to the column with an approximately flow rate of 1 ml/min. The eluat was collected in 2 ml fractions and analyzed using SDS-PAGE.

Expressed protein ligation of ER α LBD₃₀₂₋₅₂₉ α -thioester to ER α LBD₃₀₂₋₅₅₃. Native chemical ligation of the 200 μ l ER α LBD α -thioester (200 μ M) with 200 μ l peptide (4 mM) was performed in a buffer containing 20 mM HEPES/NaOH (pH 7,3), 500 mM NaCl, 1 mM EDTA, 10 % glycerol and if necessary 10 μ M E₂ in the presence of 200 μ l 150 mM MESNA (Sigma-Aldrich). The reaction was incubated for 48 h at 4 °C with slight agitation. Reaction mixtures containing a protein construct bearing a His₈-tag on the N-terminus were passed through a Ni-NTA column (HisTrap HP, 1 ml, Amersham Biosciences) to remove unligated peptide. Proteins without His₈-tag were extensively dialyzed for three days at 4°C against ligation buffer without MESNA using Slide-A-Lyzer Dialysis Cassettes (20 MWCO, Thermo Scientific) or were applied to a size exclusion column to remove the excess of peptide. Efficiency of the ligation reaction was determined by SDS gel electrophoreses and ESI mass spectroscopy.

Circular Dichroism Spectroscopy. For the CD measurements the stock solution of the ER α LBD was buffer-exchanged into 20 mM NaH₂PO₄ buffer (pH 7,7) using a Sephadex G25 PD-10 column (Amersham Biosciences). CD spectra were recorded on a Jasco J-815 CD spectrometer equipped with a peltier-type temperature controller (PTC-423S) at 10 °C using a 1 mm path length quartz cell (0.5 nm data pitch, continuous scanning mode, 100 nm x min⁻¹ scanning speed, 0.5 nm band width) at a protein concentration of 4 μ M. The measured wavelength range was from 190 nm to 250 nm. Ten scans were collected and averaged.

Mean residue ellipticity $[\theta]$ was calculated using the following equation:

$$[\theta]_R = \frac{[\theta]}{10 \times (n-1) \times c \times l} \left[\frac{\text{deg}^{-3} \times L}{\text{mol} \times \text{cm}} \right]$$

where n is the number of amide bonds, c is the concentration (mol/L), l is the pathlength (cm), $[\theta]$ is the measured ellipticity (deg⁻³) and $[\theta]_R$ is the mean residue ellipticity.

NR-coregulator interaction profiling. The On-Chip assay was performed in the absence and presence of ligand. For the preparation of the reaction mixtures purified His-ER α LBD was thawed on ice. TR-FRET Coregulator buffer E (Invitrogen, Carlsbad, USA) was complemented with DTT to a final concentration of 5 mM and all reagents were dissolved in this buffer. The final reaction mixture comprised 25 nM Alexa488-

conjugated α -His-antibody (Invitrogen, Carlsbad, USA) and 7 nM His-ER α LBD and if necessary 0.01 mM 17- β -estradiol (Serva) or *trans*-4-hydroxytamoxifen (Sigma-Aldrich). Prior use the ligand was dissolved in DMSO and diluted to a final DMSO concentration of 0.1%. Samples without ligand were also complemented with 0.1% DMSO for comparison. To confirm the sensitivity and reproducibility of the assay dilutions of the ER α proteins were performed. The dilution concentrations were 1/10, 1/50, and 1/100. All assays were performed in a PamStation-4 controlled by EvolveHT software (PamGene International BV, 's-Hertogenbosch, The Netherlands) at 20°C, at a rate of 2 cycles per minute. Nuclear Receptor PamChip Arrays (PamGene International BV, 's-Hertogenbosch, The Netherlands) contained 53 peptides on each spot. The arrays are made of a porous metaloxide carrier to which the peptides are spotted by means of Piezzo technology. Each spot has a diameter of 100 μ m and due to the porous structure the surface area is \sim 500 times larger than calculated based on spot diameter. A spot contains \sim 106 pores, each with a diameter of 0.2 μ m and a length of 60 μ m. In addition, pores are branched and interconnected. Arrays were incubated for 20 pump cycles with 25 μ l blocking (buffer 1% BSA, 0.01%, Tween-20 in Tris-buffered Saline) and then aspirated. Each array, was incubated with 25 μ l of a solution of His-ER α LBD, fluorescent α -His-antibody and if necessary ligand. Unspecific binding and antibody effects were taken into account by subtracting the simultaneous response from a reference array containing no His-ER α LBD. The solution was pumped through the porous peptide-containing membrane for 81 cycles at a rate of 2 cycles per minute and a image of every array was obtained every 20 cycles by a CCD camera based optical system integrated in the PamStation-4 instrument. For imaging a camera filter for the FITC dye with an excitation wavelength of 475 nm and an emission wavelength of 535 nm was used. Peptide microarray data analysis consisting of automated spot finding and quantitation, followed by calculation of binding velocities was performed by Bionavigator software (PamGene International BV, 's-Hertogenbosch, The Netherlands).

Cofactor recruitment FRET assay. The TR-FRET assay was performed at Organon, now Merck (Arie Visser) with 31 peptides in the absence and presence of ligand (E_2 and HT). The optimized reaction mixture contained 10 nM of purified His-ER α LBD, 1.25 nM europium-labelled α -His antibody (Eu; Perkin Elmer Life Science), 80 nM allophycocyanin-labelled Streptavidin (SA-APC; Perkin Elmer Life Science), if necessary 10 μ M E_2 and increasing concentrations of biotinylated cofactor peptide, in the assay buffer (50 mM Tris-HCl, pH 7.4, 50 mM KCl, 1 mM EDTA, 0.1 mg/ml BSA, 1 mM DTT). The dilutions of peptide ranged from 1 μ M to 0.01 nM. Each experiment was performed in duplicate in 384-well plates (Packard Optiplate) by adding 20 μ l of a receptor/ Eu-labelled α -His-antibody mixture to 10 μ l of a peptide/ SA-APC mixture followed by adding 20 μ l of 10 μ M E_2 . The reactions were mixed, centrifuged and routinely incubated overnight at 4 °C. Fluorescence at 665 nm was measured using an EnVision (2102 multilabel) Counter with a setting of 50 μ sec time delay, excitation at 340 nm, and emission at 665 nm. All data were means of two separate experiments. To directly visualize the effect of a ligand on recruitment of the different peptides the data were used for sigmoid curve plotting using Origin 7.5 (Scientific Graphing and Analysis Software, OriginLab Corp.) and the apparent EC_{50} values of peptide binding were determined.

Co-crystallization of the ER α LBD in complex with cofactor peptide and estradiol. For the crystallization the ER α LBD/ E_2 complex, purified via E_2 -sepharose, was concentrated to 12 mg/ml in a buffer containing 50 mM ammonium bicarbonate pH 7.5 and then mixed with the SRC-1 Box 2 peptide at a molecular ratio of 1.5:1 and 3:1 peptide to protein – ligand complex. For crystallization of the ER α LBD/ E_2 complex, purified via IMAC,

the concentration was performed in a buffer of 50 mM Tris-maleate pH 8.0, 50 mM NaCl, and 50 mM β -Mercaptoethanol. Before use, the peptide was dissolved in crystallization buffer to a concentration of 8.97 mM and extensively dialyzed against the same buffer for three days at 4 °C using a dialysis membrane from Spectra/Por (MWCO: 1 kDa). In order to crystallize the ER α LBD complex, initial screenings employing JCSG+, JCSG Core I, JCSG Core II, JCSG Core III and JCSG Core IV from Qiagen were performed at 20 °C and 4 °C using the sitting drop vapor diffusion. The protein concentration in the setups varied from 10 to 16 mg/ml; 0.1 μ l protein solutions were automatically mixed with 0.1 μ l reservoir solution in 96 well Corning pZero plates using a phoenix pipetting robot. The sitting drops were equilibrated against reservoirs with a volume of 70 μ l. After 2-4 days, promising thin rod-shaped crystals appeared in one condition (0.1 M Bicine, pH 9, 20% PEG 6000) at 4 °C. These crystals did not grow beyond 120 μ m x 20 μ m x 6 μ m after 21 days. However, after reproduction in hanging drops and soaking one of these crystals in reservoir solution complemented with glycerol to 20%, the crystal diffracted to about 4 Å, but was not suitable for x-ray analysis due to adhesions. In order to improve the size, morphology and quality of these crystals, streak-seeding was performed. For the preparation of the recipient drop, protein solutions, yielding crystals, were reproduced with precipitant (PEG3350) concentrations of 18% instead of 20% to avoid crystal formation prior streak-seeding. After overnight incubation at 4 °C no crystals appeared. In order to transfer the existing donor crystals (seeds) to the prepared solutions, we used a whisker and dragged the tip through the microcrystalline material. Some very small crystals remained attached to the tip of the whisker. Subsequently the whisker was streaked in a straight line across the middle of the recipient drop containing the protein and the reagent. After two days, crystals appeared along the streak line. These crystals were still highly crooked, although the final size could be improved to 300 μ m x 50 μ m x 15 μ m. Soaking of one of these crystals in reservoir solution with 20% glycerol caused a diffraction of up to 4 Å. However, due to the undesirable morphology of the crystals no x-ray analysis could be performed.

Surface Plasmon Resonance: All experiments were performed on a Biacore T100 (GE Healthcare) on CM5 chips. The running buffer used was HBS buffer (10 mM HEPES, 150 mM NaCl, 3 mM EDTA) supplemented with 0.005% surfactant P20 (Roche) at pH 7.4. For functionalizing the SH2 domain was dissolved in acetate buffer at pH 4.5 to a final concentration of 50 μ g/ml. Immobilization was performed at 25 °C using the standard EDS/NHS protocol^[27]. Subsequently, the surface was flowed with increasing concentrations of ER α LBD dissolved in HBS buffer with P20 at a flow rate of 30 μ l/min. Regeneration after each run was performed using buffer containing 10 mM glycine-HCl at pH 2.2 and 0.05% SDS. Unspecific binding and buffer effects were taken into account by subtracting the simultaneous response from a reference surface containing no SH2 domain, but functionalized with ethanolamine instead. In case of α -GST antibody (Biacore) immobilization of the membrane was incubated with a recombinant GST- Src SH2 domain fusion protein (Marligen Biosciences).

A list of all cofactor peptides used for the On-Chip studies and for the FRET studies is presented in Chapter 4.

5.10 References

- [1] a)H. Faus, B. Haendler, *Biomedicine & Pharmacotherapy* **2006**, *60*, 520; b)D. A. Lannigan, *Steroids* **2003**, *68*, 1.

- [2] N. L. Weigel, N. L. Moore, *Molecular Endocrinology* **2007**, *21*, 2311.
- [3] J. M. Wurtz, W. Bourguet, J. P. Renaud, V. Vivat, P. Chambon, D. Moras, H. Gronemeyer, *Nature Structural Biology* **1996**, *3*, 87.
- [4] a)G. Castoria, A. Migliaccio, S. Green, M. Di Domenico, P. Chambon, F. Auricchio, *Biochemistry* **1993**, *32*, 1740; b)S. F. Arnold, D. P. Vorobjeikina, A. C. Notides, *Journal of Biological Chemistry* **1995**, *270*, 30205; c)F. Barletta, C. W. Wong, C. McNally, B. S. Komm, B. Katzenellenbogen, B. J. Cheskis, *Molecular Endocrinology* **2004**, *18*, 1096; d)V. Giguere, A. Tremblay, G. B. Tremblay, *Steroids* **1998**, *63*, 335; e)G. B. Tremblay, A. Tremblay, F. Labrie, V. Giguere, *Cancer Res* **1998**, *58*, 877.
- [5] a)A. Migliaccio, G. Castoria, M. Di Domenico, A. de Falco, A. Bilancio, M. Lombardi, M. V. Barone, D. Ametrano, M. S. Zannini, C. Abbondanza, F. Auricchio, *EMBO J* **2000**, *19*, 5406; b)R. Song, H. I. Hall, R. Frey, *Stat Med* **2005**, *24*, 453; c)L. Varricchio, A. Migliaccio, G. Castoria, H. Yamaguchi, A. de Falco, M. Di Domenico, P. Giovannelli, W. Farrar, E. Appella, F. Auricchio, *Mol Cancer Res* **2007**, *5*, 1213; d)F. Auricchio, A. Migliaccio, G. Castoria, *Steroids* **2008**, *73*, 880; e)B. J. Cheskis, J. Greger, N. Cooch, C. McNally, S. McLarney, H. S. Lam, S. Rutledge, B. Mekonnen, D. Hauze, S. Nagpal, L. P. Freedman, *Steroids* **2008**, *73*, 901; f)S. B. Kim, Y. Umezawa, K. A. Kanno, H. Tao, *ACS Chem Biol* **2008**, *3*, 359.
- [6] A. M. Leduc, J. O. Trent, J. L. Wittliff, K. S. Bramlett, S. L. Briggs, N. Y. Chirgadze, Y. Wang, T. P. Burris, A. F. Spatola, *Proc Natl Acad Sci U S A* **2003**, *100*, 11273.
- [7] S. F. Arnold, J. D. Obourn, H. Jaffe, A. C. Notides, *Molecular Endocrinology* **1995**, *9*, 24.
- [8] a)P. E. Dawson, T. W. Muir, I. Clark-Lewis, S. B. Kent, *Science* **1994**, *266*, 776; b)T. M. Hackeng, J. H. Griffin, P. E. Dawson, *Proceedings of the National Academy of Sciences of the United States of America* **1999**, *96*, 10068.
- [9] T. W. Muir, D. Sondhi, P. A. Cole, *Proc Natl Acad Sci U S A* **1998**, *95*, 6705.
- [10] a)M. Q. Xu, T. C. Evans, Jr., *Methods* **2001**, *24*, 257; b)M. Q. Xu, T. C. Evans, Jr., *Methods Mol Biol* **2003**, *205*, 43.
- [11] T. Watanabe, Y. Ito, T. Yamada, M. Hashimoto, S. Sekine, H. Tanaka, *J Bacteriol* **1994**, *176*, 4465.
- [12] T. C. Evans, Jr., J. Benner, M. Q. Xu, *Protein Sci* **1998**, *7*, 2256.
- [13] a)Greenfie.N, G. D. Fasman, *Biochemistry* **1969**, *8*, 4108; b)J. P. Hennessey, W. C. Johnson, *Biochemistry* **1981**, *20*, 1085; c)S. W. Provencher, J. Glockner, *Biochemistry* **1981**, *20*, 33.
- [14] a)D. M. Heery, E. Kalkhoven, S. Hoare, M. G. Parker, *Nature* **1997**, *387*, 733; b)C. K. Glass, D. W. Rose, M. G. Rosenfeld, *Current Opinion in Cell Biology* **1997**, *9*, 222.
- [15] R. van Beuningen, H. van Damme, P. Boender, N. Bastiaensen, A. Chan, T. Kievits, *Clinical Chemistry* **2001**, *47*, 1931.
- [16] S. Folkertsma, P. I. van Noort, A. de Heer, P. Carati, R. Brandt, A. Visser, G. Vriend, J. de Vlieg, *Molecular Endocrinology* **2007**, *21*, 30.
- [17] a)A. M. Brzozowski, A. C. W. Pike, Z. Dauter, R. E. Hubbard, T. Bonn, O. Engstrom, L. Ohman, G. L. Greene, J. A. Gustafsson, M. Carlquist, *Nature* **1997**, *389*, 753; b)A. K. Shiau, D. Barstad, P. M. Loria, L. Cheng, P. J. Kushner, D. A. Agard, G. L. Greene, *Cell* **1998**, *95*, 927; c)E. H. Kong, N. Heldring, J. A. Gustafsson, E. Treuter, R. E. Hubbard, A. C. Pike, *Proc Natl Acad Sci U S A* **2005**, *102*, 3593; d)A. Warmmark, E. Treuter, J. A. Gustafsson, R. E. Hubbard, A. M. Brzozowski, A. C. W. Pike, *Journal of Biological Chemistry* **2002**, *277*, 21862; e)S. Eiler, M. Gangloff, S. Duclaud, D. Moras, M. Ruff, *Protein Expression and Purification* **2001**, *22*, 165.
- [18] U. Jonsson, L. Fagerstam, B. Ivarsson, B. Johnsson, R. Karlsson, K. Lundh, S. Lofas, B. Persson, H. Roos, I. Ronnberg, et al., *Biotechniques* **1991**, *11*, 620.
- [19] L. A. Carpino, G. Y. Han, *Journal of the American Chemical Society* **1970**, *92*, 5748.
- [20] R. B. Merrifield, *Journal of the American Chemical Society* **1963**, *85*, 2149.
- [21] a)J. C. Sheehan, G. P. Hess, *Journal of the American Chemical Society* **1955**, *77*, 1067; b)J. Hachmann, M. Lebl, *Biopolymers* **2006**, *84*, 340.
- [22] a)I. Coin, P. Schmieder, M. Bienert, M. Beyermann, *Biopolymers* **2007**, *88*, 565; b)V. Cavallaro, P. E. Thompson, M. T. W. Hearn, *Journal of Peptide Science* **2001**, *7*, 529.
- [23] K. Mullis, F. Faloona, S. Scharf, R. Saiki, G. Horn, H. Erlich, *Cold Spring Harb Symp Quant Biol* **1986**, *51 Pt 1*, 263.
- [24] G. L. Greene, C. Nolan, J. P. Engler, E. V. Jensen, *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences* **1980**, *77*, 5115.
- [25] N. S. Berrow, D. Alderton, S. Sainsbury, J. Nettleship, R. Assenberg, N. Rahman, D. I. Stuart, R. J. Owens, *Nucleic Acids Res* **2007**, *35*, e45.
- [26] J. D. Oliner, K. W. Kinzler, B. Vogelstein, *Nucleic Acids Research* **1993**, *21*, 5192.
- [27] B. Johnsson, S. Lofas, G. Lindquist, *Anal Biochem* **1991**, *198*, 268.

Chapter 6

Fragment-Based Estrogen Receptor Ligands

Part of this work has been published: S. Renner, W. A. van Otterlo, M. Dominguez Seoane, S. Möcklinghoff, B. Hofmann, S. Wetzel, A. Schuffenhauer, P. Ertl, T.I. Oprea, D. Steinhilber, L. Brunsveld, D. Rauh, H. Waldmann, *Nat. Chem. Biol.* **2009**, 5(8):585-92

Abstract: Based on a new scaffold, identified via bio-activity guided mapping, novel fragment-based agonists for the estrogen receptor were designed and evaluated via biochemical, cell-biological and structural studies. A focused fragment library enabled the establishment of a clear Structural – Activity - Relationship (SAR).

6.1 Introduction

The estrogen receptor (ER) belongs to the superfamily of nuclear receptors (NRs)^[1]. NRs are multidomain transcription factors typically under the control of a ligand – a small lipophilic molecule that easily penetrates biological membranes – that shifts the equilibrium of the conformation of the NR between an agonistic and an antagonistic state^[2]. Binding of an agonist to the ER ligand binding domain (LBD) stimulates the interaction of the NR with coactivators, resulting in transcriptional activity. In contrast, the complex of ER LBD and an antagonist recruits corepressors that typically repress transcriptional activity^[3]. The primary endogenous ligand for the ER is the estrogen 17 β -estradiol (E₂; **1**)^[2]. E₂ is present in circulation from the onset of puberty to menopause and can lead to breast cancer and troubles in menopausal women.

Blocking and/or preventing breast cancer is the major therapeutic role of ER antagonists such as tamoxifen and its active metabolite 4-hydroxytamoxifen (HT; **3**)^[4]. Symptoms in menopausal women can be treated using ER agonist-directed therapy, *e.g.* through the use of hormone replacement therapy (HRT) using the agonist diethylstilbestrol (DES; **2**; Figure 1)^[5]. However, current estrogen-based therapies have serious side-effects in non target tissues that present serious challenges^[6]. Understanding the exact molecular interactions of synthetic ligands with specific NRs is an essential step in identifying novel tissue-selective synthetic modulators that minimize unwanted side activities.

Two subtypes of the ER (ER α ^[7] and ER β ^[8]) with unique tissue distribution patterns and transcriptional properties are known. While ER α is mainly involved in reproduction events in uterus and mammary gland^[9], ER β is more generally expressed and not the dominant receptor in uterus and breast tissues. The challenge to identify highly ER β selective modulators provides the entry to elucidate the exact physiological function of ER β and to develop potential novel, tissue and cell-selective drug candidates such as those related to inflammatory diseases (Figure 1)^[4b, 10].

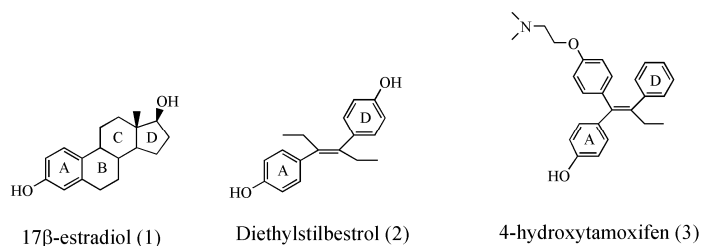


Figure 1: Chemical structures of the natural ER estrogen 17 β -estradiol (**1**), the synthetic estrogen diethylstilbestrol (**2**) and the synthetic anti-estrogen 4-hydroxytamoxifen (**3**).

Fragmentation of bioactive molecules into discrete functional groups can be used to simplify the analysis of ligand binding and to map out different elements required for high affinity. Due to the flexible adjustment of the small fragments, this approach allows the simple optimization of unique interactions in the binding site to develop compounds with improved binding affinities.

This chapter describes the design of ER modulators based on a simplified fragment scaffold for which ER activity has not been annotated before. Incorporation of variable groups with respect to the structure – activity relationship (SAR) and fragment-based optimization were further used to increase the affinity of these compounds for both receptors ER α and ER β . X-ray crystallographic studies of these compounds complexed with the ER β LBD helped to further understand the mechanism of ER β affinity and selectivity and provide the entry to develop new ER β modulators with improved affinity and selectivity.

6.2 Identification, Design and Fragment-based Synthesis of Novel Estrogen Receptor Ligands

Current fragment-based screening and design approaches utilize small molecules with low affinity as starting points for the synthesis of larger molecules with higher affinities and selectivities. In an effort to identify novel non-steroidal ER modulators we focused on the reverse search, the identification of smaller structural simplified scaffolds incorporated into larger known molecules with which they likely share biological properties. Structural simplification may guide the design of synthetically more tractable molecules retaining elements of the biological relevance of the complex molecules^[11]. An *in silico* compound library, in which the molecules are chemical organized in tree-like branches from structural complex to more simple scaffolds, was searched. The compounds within a given branch of the scaffold tree are not only structurally related, but may also share bioactivity on the same biological target^[12]. Using this structural minimization approach tetrahydroisoquinoline (THIQ; **4**) was identified as a simple chemical 2-ring core scaffold for which no ER activity has been reported before^[12]. THIQ actually represents a simplified mimic of estradiol, in which the C and D rings of estradiol have been removed, providing us with a further understanding of their role in ER binding and activity. Incorporation of a nitrogen atom in the B ring allows for the rapid synthesis of substituted analogues and enables THIQ to be a suitable core scaffold for the design of novel ER modulators. Indeed, complex THIQ derivatives with at least three ring scaffolds have been described previously as selective ER modulators^[13]. Some selected examples for THIQ analogues as ER modulators are shown in

Figure 2. Generally these compounds resemble the structure of the SERM lasofoxifene (**5**)^[4b, 14], a drug used for the treatment of osteoporosis. Compounds with bulky side-chains, mimicking the phenyl alkoxyamine moiety found in lasofoxifene, could be shown to act as antagonists in mammalian cell studies and mostly feature an advanced selectivity for the ER α (**6-7**)^[13c, 13d]. Similar binding affinities could be observed, when the N-substituent in **6** was exchanged by a phenol or naphthyl group, connected via a sulfonamide linker^[13c]. Less THIQ analogues could be identified so far bearing agonistic characteristics and an advanced selectivity for the ER β . These compounds typically feature a less bulky or no side-chain at the C₁ position in the THIQ scaffold (**8-9**)^[13c, 15].

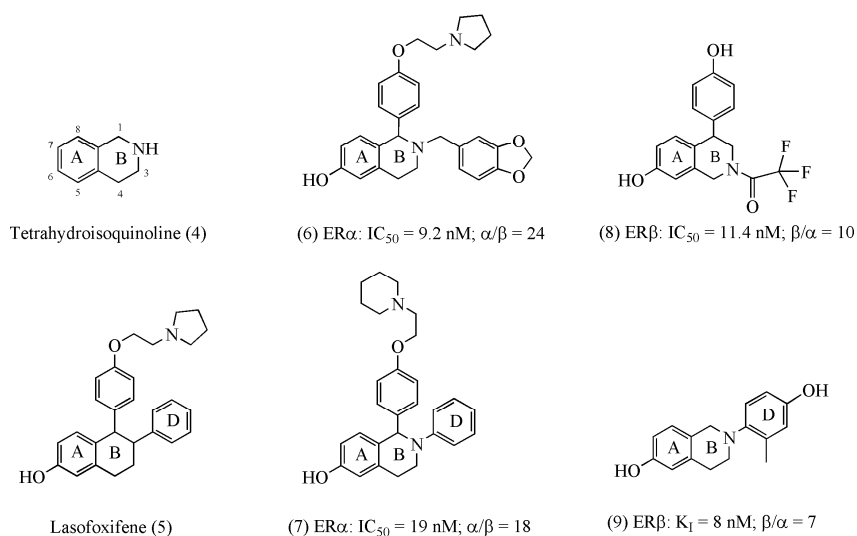


Figure 2: Chemical structures of tetrahydroisoquinoline (**4**) and several synthetic ER modulators derived from the SERM lasofoxifene (**5**) with binding affinities and selectivities for an ER isoform.

Based on the simple tetrahydroisoquinoline scaffold (**4**) several related molecules were designed and investigated for their binding affinities for both receptors the ER α and ER β . Comparison of the THIQ with the structure of the natural ER agonist estradiol (E₂) indicates that the two rings most probably mimic the A- and B-rings of E₂ and seems to be an essential requirement for synthetic ER ligands to possess good binding affinity. Therefore it was decided to introduce an OH-group at the C₆ position to mimic the phenolic group in the A-ring of E₂ which is known to be essential for its binding to the ER. Moreover, incorporation of bulky side chains as N-substituents of the THIQ scaffold was thought to fill the space

normally occupied by the C- and D-rings of E₂ and thus may provide the greatest opportunity for further enhancing ER affinity. A variety of functional groups was examined including electronegative, aromatic and polar N-substitutions. Starting from the commercial available 1,2,3,4-tetrahydroisoquinoline, 22 THIQ derivatives were synthesized by Willem van Otterlo using standard chemical transformations. Synthesis of THIQ derivatives, including the phenolic group at the C₆ position, started from the commercial N-protected tetrahydro-6-isoquinolinol according to a strategy published by Hoye *et al.*^[16], which utilized an *in situ* protection of the phenolic group with a TES group. The purity and integrity of the THIQ analogues was confirmed by NMR studies and if possible compared to previously published data. A summary of all synthesized compounds is shown in Table 1.

6.3 *In vitro* Structure – Activity Relationship of Novel Estrogen Receptor Ligands

The newly synthesized THIQ analogues were evaluated in a biochemical fluorescence polarization (FP)^[17] assay. In this assay, the individual binding affinities (EC₅₀) of the potential ER modulators for the ER α and ER β LBD were measured indirectly by means of the recruitment of a fluorescently labeled cofactor peptide, derived from the coactivator SRC 1 box 2. This molecular probe binds to the agonist-liganded state of the ER LBD via its LXXXLL recognition motif.

For the preparation of the protein moiety, the ER α and ER β LBD were separately expressed in *E. coli* with an N-terminal His-tag in the absence of any ligand. After purification of the apo-proteins via immobilized metal ion exchange chromatography (IMAC), their purity was confirmed by SDS-PAGE^[18]. The cofactor peptide was synthesized via standard solid-phase peptide synthesis (SPPS)^[19] with an N-terminal cysteine and labeled with fluorescein. For the ligand binding studies, the fluorescein labeled peptide (fl-peptide) was combined with a 40-fold excess with either of the ER LBDs and titrated with increasing concentrations of each compound, dissolved in assay buffer including 5% DMSO. Each measurement was performed in triplet and solutions of buffer with 5% DMSO were used as negative controls.

FP^[17] was used to detect changes in the size of the molecular complex by monitoring the polarization state of the fluorescence output when excited with linear-polarized light. In absence of ligand, the small fl-peptide is not able to bind stably to the ER LBD and rotates rapidly in solution resulting in a depolarization of the fluorescence signal. Interaction of the small molecules with the ER LBD can cause various effects. Binding of the compound as an agonist, will result in the formation of a complex of the ER LBD and the fl-peptide, which

rotates more slowly maintaining the exposed light polarized (Figure 3). Antagonistic ER modulators would totally repress the binding of the fl-peptide to the ER LBD resulting in a further decrease in the fluorescence signal.

The calculated effector concentrations for half-maximum response (EC_{50}) values of all compounds are presented in Table 1. As expected, the endogenous ligand E_2 binds in the same range to both ER isoforms in these *in vitro* studies. An initial comparison of the *in vitro* data revealed that 12 out of the 22 synthesized THIQ analogues showed activity for both ER LBDs. These 12 compounds all featured an agonistic profile, including the interaction of the ER LBDs with the coactivator peptide (Figure 3).

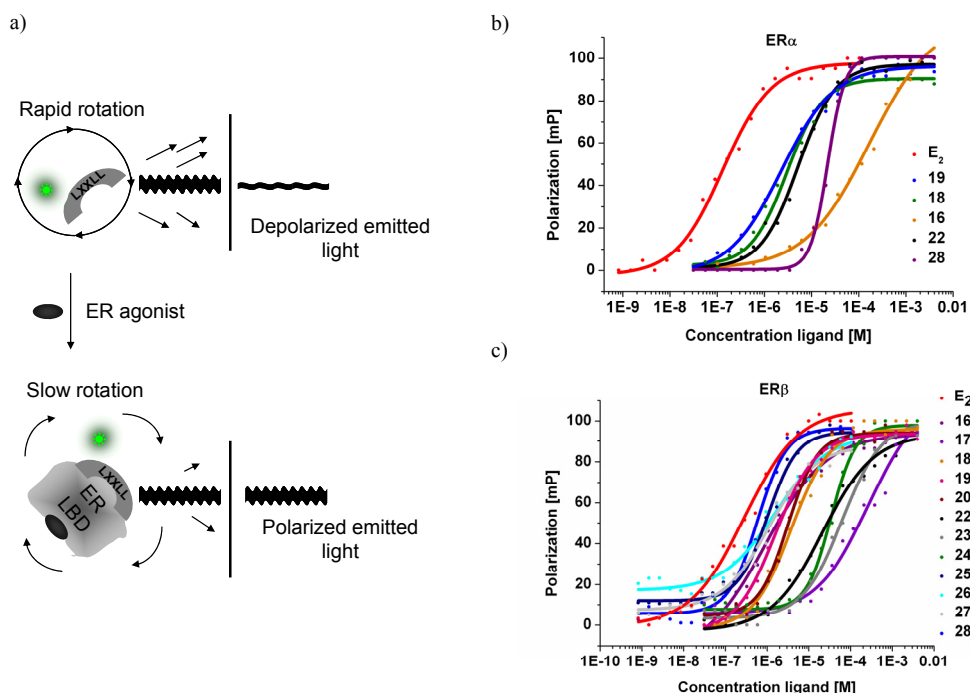
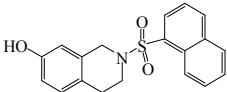
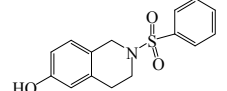
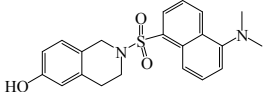
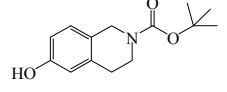
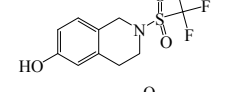
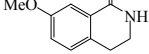
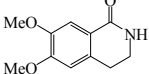


Figure 3: (a) Schematic representation of the fluorescence polarization assay. A fluorescein labeled cofactor peptide with high initial fluorescence depolarization allows to evaluate the binding of those substances that demonstrate affinity for the ER, thereby changing the polarization of the peptide and to assess the selectivity of the compounds toward the ER α (b) or the ER β LBD (c).

Table 1: Estrogen receptor binding affinities (EC₅₀) of simple 2-ring tetrahydroisoquinoline derivatives

Compound	Structure	Binding affinity (μM) ^a	
		ERα LBD	ERβ LBD
1		0.134	0.28
10		n.a.	n.a.
11		n.a.	n.a.
12		n.a.	n.a.
13		n.a.	n.a.
14		n.a.	n.a.
15		n.a.	n.a.
16		5.22	0.9
17		340	227
18		20	3.77
19		2.33	1.6
20		6.2	3.17
21		n.a.	n.a.
22		170	20
23		700	60

24		600	30
25		Not tested	1.03
26		Not tested	1.59
27		Not tested	1.09
28		3.02	0.59
29		n.a.	n.a.
30		n.a.	n.a.

n.a. no activity, ^a Values are means of a least three experiments

Initially it was chosen to keep the THIQ scaffold as simple as possible and only explore the space described by various N-substitutions. Incorporation of side-chains at this position was thought to fill the space normally occupied by the C and D-ring of estradiol. As expected, the unmodified THIQ scaffold did not feature any binding for the ER LBDs in our FP studies. Previous studies indicated that the incorporation of an OH-group mimicking the phenolic A-ring of estradiol enhances the binding affinity for the ER^[20]. Incorporation of either a 6-hydroxyl group (**10**) or a 7-hydroxyl group (**11**) revealed that the OH-group alone is not able to achieve ER binding. Also the introduction of both OH-groups at the same time could not facilitate ER potency (**12**). Next, it was focused on various N-substitutions. Attachment of just a simple acetamide did not allow compound **13** to interact with the ERs even not when it features a 6-hydroxyl group (**14**). Further analysis revealed that only N-substitution with somewhat bulky and electron withdrawing groups like the trifluoroacetamide group improves the affinity for both ER isoforms. However, N-substitution alone did not facilitate ER affinity as shown for **15**. Only the concomitant incorporation of a 6-OH-group caused a detectable affinity for the ER LBDs (**16**). Masking of the phenol using a methoxy protection group in **17** resulted in a 200-fold decrease in binding affinity for both ER LBDs. Exchanging the 6-hydroxy group by a 7-hydroxy group (**18**) causes a 4-fold loss in ER α potency and a 7-fold

loss in ER β potency. These results revealed that the availability and the position of the OH-group at the C₆ atom are essential for high affinity binders. Beside the phenolic hydroxyl group of the A-ring, the composition of the N-substitution was crucial for ER binding. The replacement of one or all three fluoro groups in **16** by chloro groups was also well tolerated (**19** and **20**). However, a small decrease in binding affinity could be observed with increasing numbers of fluoro to chloro exchanges. This effect can be attributed to the lower electron negativity of the chlorides, relatively to the fluorides. The less electronegative chloride most probably reduces the polarity of the THIQ core to less extent than the highly electronegative fluoride. Alternatively, the larger size of the chloro groups might play a role as well.

An additional set of compounds, bearing a sulfonamide group as N-substitution (**21**), was investigated. Similar to the trifluoroamide group these THIQ derivatives lead to ER binders, in combination with a 6-hydroxy group (**22**). Enlargement of the side-chain by incorporation of a 1-naphthyl group sulfonamide led to the less potent compound **23**. Interestingly, analog **23** showed a 10-fold preference for ER β over ER α . With this, compound **23** features the highest selectivity for ER β of all THIQ analogues measured in these *in vitro* studies. The 7-hydroxy THIQ analog **24** does not decrease the ER potency. It has to be noticed that **23** and **24** were badly soluble in the reaction buffer, indicating a possibly better EC₅₀ than measured in the current set-up. Reduction of the steric bulk by replacing the 1-naphthyl group by a phenyl ring (**25**) resulted in a 35-fold increase in ER β affinity. Introduction of a dimethylamino analog of the naphthyl group (**26**) featured an almost equal potency for ER β as **25** indicating that a reasonable steric bulk could be tolerated if appropriately positioned. Possibly the dimethylamino hydrogen bond acceptor unit plays an important role here. Apparently, the ER β binding affinity is improved by slightly bulky substituents. For example, incorporation of a 2-carboxylic acid *tert*-butyl ester (**27**) exhibits an ER β binding affinity similar to **25** and **26**. The importance of electronegative functionalities, also for the sulfonamide series, is highlighted by compound **28**, featuring a sulfonamide with an electronegative trifluoromethane group. This compound has a notable ER potency and is the best ER binder in this series and represents reasonably good affinity for the ER β (EC₅₀: 0.59 μ M) with an interesting ER β selectivity (β/α : 5). Finally, compound **29** and **30** were designed in order to evaluate the effect of an oxygen substitution in position C₁. To guarantee that a possible observed effect can not be attributed to an additional modification, a 7-methoxy group (**29**) and an additional 6-methoxy group (**30**) was used. Further, no N-substitution was incorporated. The complete lack of activity for both ER isoforms suggests that a simple oxygen group in position C₁ is not sufficient to enable ER binding affinity.

Comparing the EC₅₀ values of the THIQ analogues it is striking that all compounds feature selectivity for the ER β LBD. The general binding affinities of the best ER β LBD binders (**28** > **16** > **27** > **25** > **26** \geq **19**) were determined to be as potent as estradiol in this FP studies. In contrast, the best determined ER α LBD binder was the THIQ derivate **19** (EC₅₀: 2.33 μ M), which featured a 23-fold weaker potency then the endogenous substrate estradiol. It should be noted that due to the high protein concentration (4 μ M) used in this FP study, the determined EC₅₀ values do not represent the K_D, as it is for example known that the K_D of estradiol lies in the sub nanomolar regime.

6.4 Transactivation efficiency of the novel Estrogen Receptor Ligands *in vivo*

In order to study the transcriptional activity of the new fragments on ER α and ER β in mammalian cells, E₂ and three selected THIQ analogues were profiled in a cellular transcription assay using U2SO cells^[21]. U2OS human osteosarcoma cell lines do not endogenously express ERs at a detectable level. However, transient transfection of these cells with mammalian expression constructs for either the human ER α or the ER β full-length sequences result in the expression of the receptors. The transactivation capacity of transfected ERs was confirmed by an estrogen response element (ERE) – firefly luciferase (LUC) reporter gene assays. In this assay, agonists like estradiol induce the ERE reporter gene activity by recruiting the ER to the ERE and allowing the interaction with coactivators and subsequent DNA polymerase resulting in the transcription of the downstream luciferase gene (Figure 4). The expressed firefly luciferase^[22] catalyses the monooxygenation of beetle luciferin, a process emitting photons. In order to minimize experimental variability caused by differences in cell viability or transfection efficiency, the mammalian cells were co-transfected with a second reporter gene expressing the *Renilla* luciferase^[23] and the activity of the experimental reporter was normalized to the activity of this internal control. This cellular luciferase assay allows a more detailed determination of the exact binding affinities for the ERs and an evaluation of the capacities of cellular activity and selectivity of the molecules.

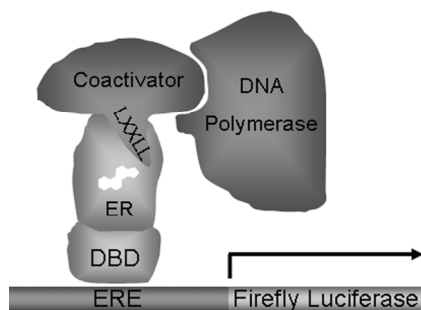


Figure 4: Mechanism of the estrogen induced luciferase transcription. Binding of estrogen to the LBD allows the ER to interact with coactivators via the LXXLL motif resulting in ERE binding upstream of the gene for the firefly luciferase. Recruitment of the DNA polymerase by the ER – coactivator complex activates gene transcription followed by expression of the luciferase.

Comparison of the general transfection efficiency and transcriptional activity of both ERs was performed by incubating the transfected cells in the absence of ligand and in the presence of a fixed concentration of E_2 and the THIQ analog **28** (Figure 5). In agreement with previous studies^[24], E_2 caused a 33-fold increase in luciferase activity in $ER\alpha$ -transfected cells, which indicates a high level of receptor activity. In the case of $ER\beta$ -transfection, E_2 -treatment resulted in an only 7-fold increase in luciferase activity thus indicating a moderate, but sufficient level of receptor activity. However, exposure of the cells to the same concentration of compound **28** (10 nM) could neither increase the basal luciferase expression in $ER\alpha$ - nor in $ER\beta$ -transfected cells (Figure 5).

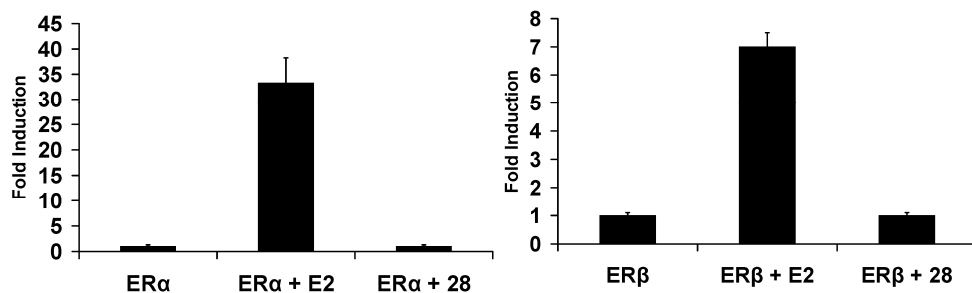


Figure 5: Confirmation of transactivation of transfected $ER\alpha$ and $ER\beta$ by ERE-Luciferase reporter gene assays. U2OS cells were transfected with a LUC reporter plasmid under the control of an ERE and a human $ER\alpha$ or $ER\beta$ plasmid. The luciferase activity was normalized to the *Renilla* activity and set to 1 in the absence of ligand. Shown is one representative of three independent measurements. Compound **28** featured no Luc-activity at the given concentration (10 nM ligand).

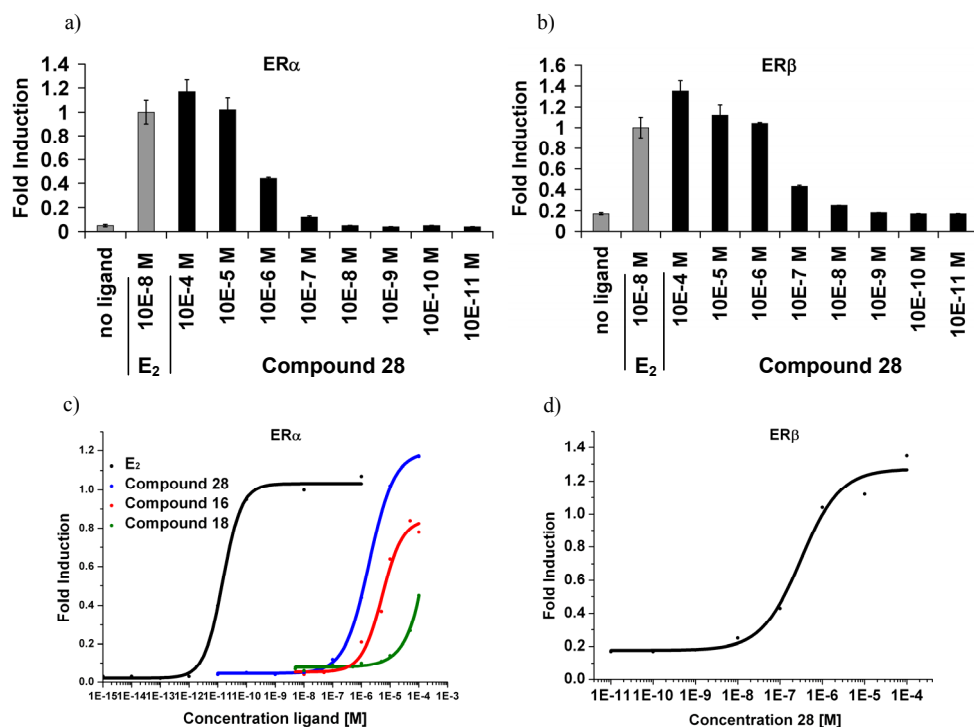


Figure 6: Concentration dependent luciferase induction of compound **28** in ER α -transfected U2OS cells (a) and in ER β -transfected U2OS cells (b) referred to the normalized luciferase activity in the presence of 10 nM E₂. (c) Logarithmic plotting of the normalized luciferase activity allowed the determination of the binding affinity (EC₅₀) of E₂ (black), compound **16** (red), compound **18** (green) and compound **28** (blue) for ER α . The binding affinity of the best THIQ binder was approximately 5 magnitudes lower than estradiol. (d) For ER β the EC₅₀ of only compound **28** could be determined.

Table 2: Estrogen receptor binding affinities (EC₅₀) of a selected set of THIQ analogues by means of a cell-based luciferase transactivation assay.

Ligand	ER α [M]	ER β [M]
Estradiol	1.49 × 10 ⁻¹¹	around 1.00 × 10 ⁻¹⁴
Compound 16	5.3 × 10 ⁻⁶	Not tested
Compound 18	8.52 × 10 ⁻³	Not tested
Compound 28	1.87 × 10 ⁻⁶	3.11 × 10 ⁻⁷

At a concentration of 100 nM and higher, an increase in luciferase activity could be detected for compound **28** (Figure 6 a-b) for both ERs isotypes. For a better comparison of the binding affinities of estradiol and several THIQ analogues for the ERs, dose response

analyses were performed for E₂ and a selected set of THIQ derivatives. Plotting of the measured luciferase activities against the logarithmic ligand concentrations allowed the exact determination of the binding affinities (EC₅₀) of the designed compounds relative to E₂ (Figure 6 c-d and Table 2). In agreement with the *in vitro* studies, all compounds show agonist activity, with compound **28** featuring the highest affinity for ER α , with an EC₅₀ value of 1.87 μ M, followed by compound **16** (EC₅₀: 5.3 μ M) and finally the moderate ER α binder **18** (EC₅₀: 85.2 μ M). Nevertheless, the binding affinity of E₂ was five magnitudes higher (EC₅₀: 14.9 pM) than measured in the *in vitro* FP assay (EC₅₀: 0.13 μ M). For ER β the luciferase activity for E₂ and compound **16** and **18** could not be determined due to low transactivation signals. The EC₅₀ of **28** for ER β (0.31 μ M) is rather potent for such a small fragment almost tenfold more potent than on ER α . The results from the transactivation studies indicate that the designed THIQ derivatives bind the ERs in the micromolar regime and are active on the cellular level. The compounds furthermore show an interesting ER β selectivity and thus could constitute good hits for further optimization. In particular, the affinity of **28** is notable with respect to the size of the ligand. Moreover, compound **28** features the highest measured ER β selectivity in these studies.

6.5 Co-Crystallization of Novel Estrogen Receptor β Ligands

To elucidate the binding affinity and selectivity of the fragments, structural information was obtained on the ligand-ER β LBD complexes. X-ray crystallographic studies^[25] additionally could be valuable for the design of follow-up libraries. For this, the ER β ligand binding domain (LBD) was expressed in *E. coli* in the absence of ligand and subsequently purified via estradiol affinity chromatography. The bound ER β LBD could be eluted from the immobilized estradiol-column via displacement with an excess of the desired ligand. The ligand-occupied ER β LBDs were further purified via size exclusion chromatography (SEC) and the purity of the protein was subsequently confirmed by SDS-PAGE (Figure 7). For the X-ray studies it was decided to compare the crystal structure of the ER β LBD bound to the natural agonist estradiol with the best THIQ scaffolds in this series, **28** and **16**. In order to investigate the influence of the position of the hydroxyl group of the phenolic ring, the complex of ER β LBD and the moderate binder **18** was chosen as well for co-crystallization.

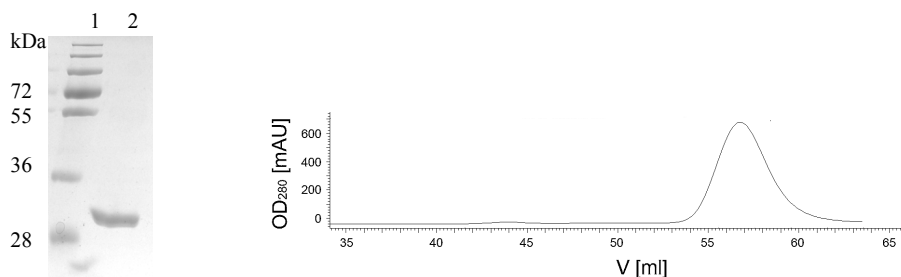


Figure 7: Analysis of the purification of the ER β LBD. (a) SDS-PAGE gel (15%; Coomassie stained). Lane 1: Molecular weight marker, lane 2: ER β LBD_{MD[D261-L500]DD} (calcd. mass: 27553.5 Da). (b) Elution profile of ER β LBD_{MD[D261-L500]DD} after SEC using FPLC.

The concentrated ligand-bound ER β LBD was incubated with a synthesized coactivator peptide derived from the steroid receptor coactivator 1 (Src1 Box 2 ⁶⁸⁵LTERHKILHRLLEQEGSPSD⁷⁰³; calculated mass: 2229.52 Da, detected mass: 2228.54 Da). Primarily, the ER β LBD complexed with the endogenous ligand estradiol was screened for suitable crystallization conditions. Optimal crystals were grown at 4 °C from a hanging drop containing a mixture of protein-estradiol solution and reservoir solution of 17% PEG 10000 (v/v), 0.1 M Ammonium Acetate and 0.1 M Bis-Tris pH 5.5 (Figure 8). These crystals reached a final size of 0.5 x 0.4 x 0.3 mm within seven days. Prior to data collection the crystals were briefly soaked in a solution containing the mother liquor and 20% glycerol (v/v) and flash-frozen in liquid nitrogen.



Figure 8: Co-Crystals of the ER β LBD in complex with cofactor peptide and the variable ligands E₂, **28**, **16** or **18** under conditions of 0.2 M Ammonium Acetate, 0.1 M BIS-TRIS pH 5.5 and 17% PEG10000 (v/v) at 4 °C. The overall size of the crystals is 0.5 x 0.4 x 0.3 mm.

The overall crystal structure of the E₂-ER β LBD complex was similar to those previously reported^[26] (Figure 9). The ligand was completely excluded from the external environment by the position of helix 12 (H12), which folds over the binding pocket and projects the hydrophobic side chains of the ER β residues Leu₄₉₁ and Leu₄₉₅ towards the bound

ligand. The conformational change of H12 allows the formation of a hydrophobic groove on the surface of the ER β LBD, created by the helices H3, H4, H5 and H12^[26c]. The coactivator peptide occupies this cavity via the LXXLL motif in an α -helical conformation^[27] (Figure 9a).

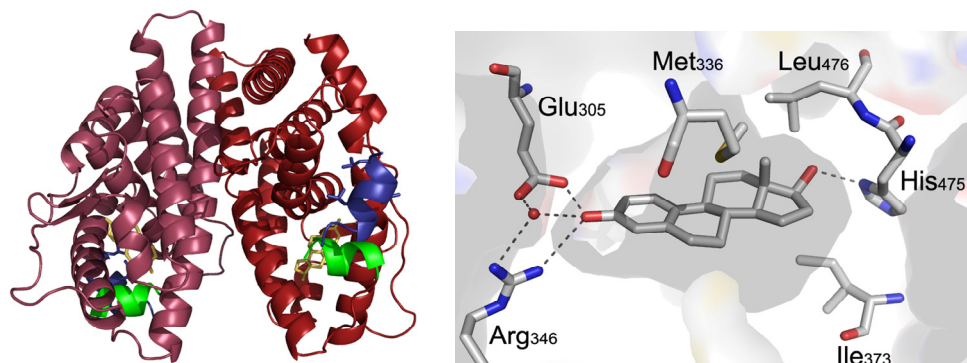


Figure 9: Left: The three-dimensional protein structure of the dimeric E₂-ER β LBD complex including a coactivator peptide (blue) bound to the hydrophobic groove on the surface of the ER β LBD; E₂ is shown as a stick model (yellow) and helix 12 is presented in green. Right: The interaction of E₂ with critical amino acids in the ligand binding pocket of ER β LBD; only key residues are shown for simplicity and hydrogen bonds are presented as dotted lines.

E₂ binds diagonally across the deeply buried ligand binding pocket between H11, H3, and H6 and adopts a low-energy conformation. A combination of specific hydrogen bonds and van-der-Waals contacts of the binding cavity allow the high affinity binding of the non-polar E₂. In particular, the phenolic hydroxyl group of the A-ring is essential for the correct anchoring of the ligand. This OH-3 forms a hydrogen bond to Glu₃₀₅ from H3 and to Arg₃₄₆ from H6 and a highly ordered water molecule^[20] (Figure 9 b). A second hydrogen bond is formed between the 17 β -OH in the D-ring of estradiol and His₄₇₅ in H11 at the distal end of the cavity^[26c]. However, early studies suggested that the 3-OH group of the A-ring of E₂ contributes an average of 1.9 kcal/mol to the free energy compared to only 0.6 kcal/mol for the 17 β -OH in the D-ring^[28]. Therefore, the ability to mimic the A-ring, present in the natural estrogen E₂, seems to be the most essential requirement for synthetic ER ligands to possess good binding affinity. In agreement with this, the binding affinity of the novel designed THIQ analogues for ER drastically increased if an OH-group was attached to the correct position on the A-ring.

In order to define the exact binding mode of the THIQ derivatives, co-crystallization studies were performed using the ER β LBD complexed with the two best binders in this

study, compound **28** and **16**. Further, the moderate binder compound **18** was chosen for X-ray studies. Expression and purification of the ligand-occupied ER β LBD were carried out like described above for the E₂-bound ER β LBD. For the crystal screening the same conditions as for estradiol were chosen and optimal crystals appeared after one day at 4 °C, reaching a final size of 3 x 2 x 1 mm within the next seven days (Figure 8). Shape, size and crystallization behavior of the different crystals was similar to the crystallization of the E₂-ER β LBD complex. These crystals diffracted to at least 2.2 Å and were suitable for X-ray analysis. Moreover, the overall solved structures of the ER β LBD complexed with **28**, **16** and **18** are similar to those of the E₂-ER β LBD complex^[26] and thus they will not be described in detail here (Figure 10), but rather the focus will be on the ligand bound in the cavity. It has to be pointed out that in all three cases the helix 12 conformation corresponds to that observed for E₂ and other bound agonists, consistent with the observed functional agonistic activity derived from the *in vitro* FP and the *in vivo* transactivation studies.

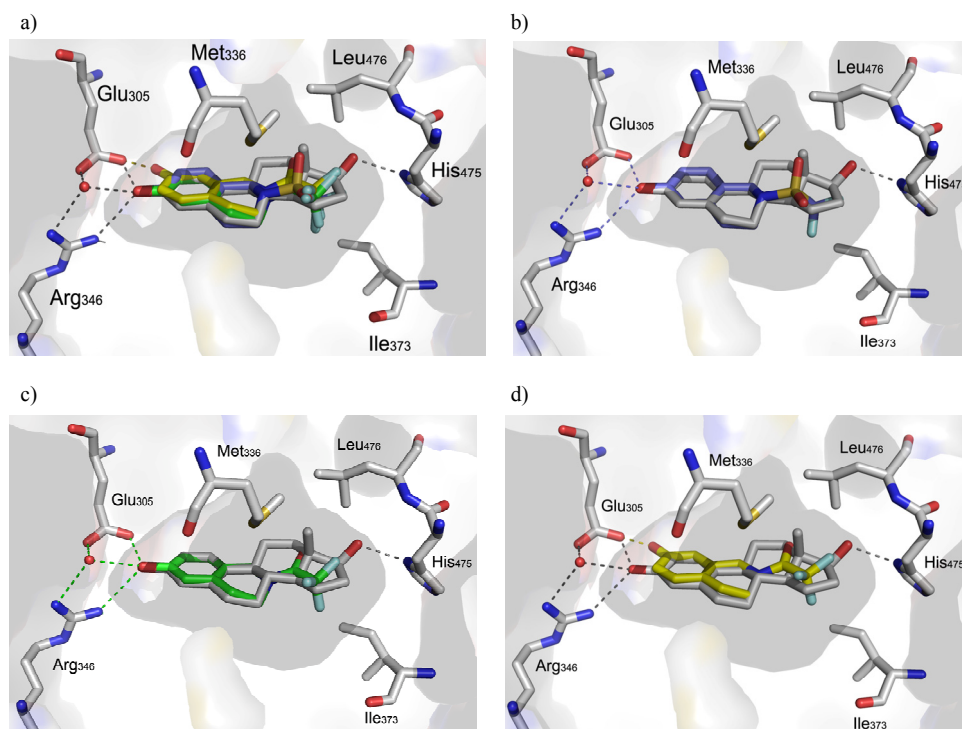


Figure 10: (a) Overlay of the X-ray structures of the ER β LBD co-crystallized with E₂, compound **28**, compound **16** and compound **18**; (b) ER β LBD-**28** (blue) overlaid with ER β LBD-E₂ (grey), (c) ER β LBD-**16** (green) overlaid with ER β LBD-E₂ (grey) and (d) ER β LBD-**18** (yellow) overlaid with ER β LBD-E₂ (grey). Only key residues are shown for simplicity. Hydrogen bonds to key residues are shown as dotted lines.

Unbiased electron density difference maps unambiguously define the ligand binding mode of all complexes studied. An overlay of the X-ray structures of the ER β LBDs bound to **28**, **16** and **18** indicates that they adopt similar positions (Figure 10a). For all three ligands, the binding modes of the two planar ring systems were identical to the A-ring and the B-ring of E₂ (Figure 10b-d). In particular, the phenolic hydroxyl of the compounds **28** and **16**, having their hydroxyl groups in comparable positions to that of estradiol, is involved in the above mentioned Glu₃₀₅-Arg₃₄₆-HOH hydrogen bonding triad (Figure 10b-c). These compounds were also observed to have the highest affinities for ER β in this series. Comparing the crystal structure of compound **18** it can be noticed, that the 7-hydroxyl group forms only one hydrogen bond with Glu₃₀₅, but neither with Arg₃₄₆ nor with the water molecule (Figure 10d). In agreement with this observation, the ER β binding affinity of compound **18** is 7- and 9-fold decreased versus **28** and **16**, indicating that the position of the OH group in the phenol plays a significant role in determining ER affinity. The backbone of all three compounds established a number of van der Waals contacts with the residues forming the binding pocket.

While the rigid protein architecture around the A-ring pocket imposes an absolute requirement on effective ER ligands to contain a planar ring group, the remainder of the binding cavity is quite accommodating^[27a]. In particular, the distal end of the cavity is quite flexible and permits a variety of ligand-binding modes^[26a-c]. The N-substitutions of the three crystallized THIQ derivatives occupy a position similar to that adopted by the C- and D-rings of estradiol (Figure 10b-d). However, due to the lack of a second OH-group mimicking the 17 β -OH in the D-ring of estradiol, the THIQ analogues are not able to form a hydrogen bond with the His₄₇₅ side chain at the distal end of the ER β cavity (Figure 10b-d). This interaction is known to further stabilize the ligand – LBD complex^[26c]. In fact, the binding affinity of all designed compounds is around five magnitudes lower, relatively to estradiol. However, the interaction of the ligand via hydrogen bonds contributes only in part to the affinity of the ligand for the ER β . The performed binding studies indicate that special N-substitutions of the THIQ derivatives improve the ER β potency. In particular, the combination of bulky and electron-withdrawing groups, as realized for compound **28** and **16**, turned out to be most suitable. The affinity enhancement for both ER isoforms after incorporation of halogen groups can be attributed to a favorable overall hydrophobic effect due to the bulky substituent itself as well as to reduced polarity of the THIQ core induced by the electronegative halogens. Additionally, van der Waals interactions between the halogen and surrounding residues further stabilize the ligand within the ligand binding pocket.

Significantly, compounds **28** and **16** featured not only the best ER potency of this series, but also displayed a 4- and 5.5-fold ER β selectivity *in vitro* and an even slightly higher selectivity in cellular assays. Although the compounds **23** and **24** displayed an even better selectivity (10-fold) for ER β , their binding affinity for the ER was almost 10-fold weaker compared with **28** and **16**. In general, the smaller overall binding pocket of ER β relative to ER α would favor smaller molecular structures^[26a]. The observed selectivity for the bulky compounds **23** and **24**, however, is consistent with the specific observation that aromatic moieties appear capable of making a more favorable interaction with amino acids in the cavity of ER β than with residues in the ER α ^[26b]. Responsible for the difference in the size of the cavity are two amino acid residues in close proximity to bound agonists: ER α Leu₃₈₄ is replaced by ER β Met₃₃₆, and ER α Met₄₂₁ is replaced by ER β Ile₃₇₃^[26, 27b, 29]. Although ER α and ER β are only 58% identical in sequence, these two amino acid residues are the only two conservative residue changes of the two ER isoforms within the ligand binding pocket^[26a]. Ligands that are able to interact differently with ER β Ile₃₇₃ than with ER α Met₄₂₁ have been proposed to enhance the selectivity for ER β ^[26d, 29-30]. Briefly, the substitution of Met₄₂₁ by Ile₃₇₃ may impart the ability of certain appropriately placed functional groups to achieve stereoelectronic differentiation between these two residue side chains, leading to enhanced ER β selectivity. As shown in Figure 11, the trifluoroacetamide moiety of the crystallized THIQ derivatives is clearly in close proximity to the Met₄₂₁/Ile₃₇₃ residues. In particular, it appears likely that the low distance between one of the fluoro atoms and the sulfur of ER α Met₄₂₁ represents a repulsive interaction, which may be limiting the conformational space explored by the methionine side chain. Although the relative contribution of dispersion, electrostatics, and exchange repulsion is unclear, it is possible that the electronegativity of the halogens and the methionine sulfur makes an unfavorable electrostatic contribution to the total interaction^[26d]. In contrast, the distance between the electron-withdrawing halogen atom and ER β Ile₃₇₃ is greater, which is not expected to result in an unfavorable interaction (Figure 11). This observation is consistent with the measured ER β selectivity of compounds **28** and **16**.

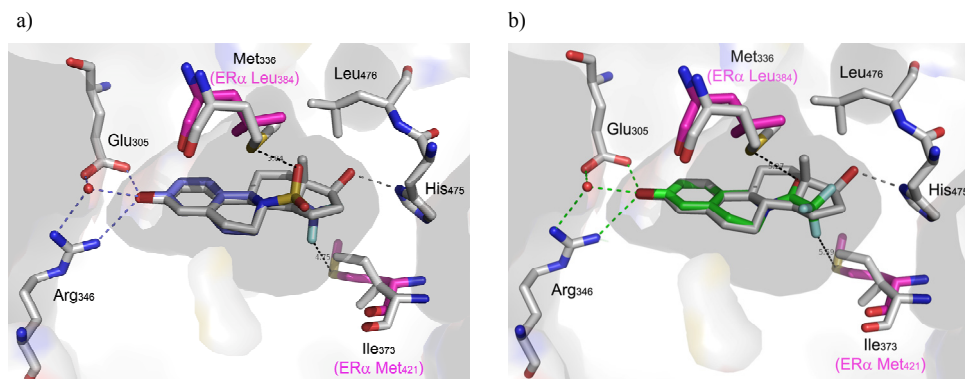


Figure 11: The crystal structure of compound **28** (a) and compound **16** (b) complexed with the ER β LBD (grey) and overlaid with ER α LBD – E₂ (magenta; PBD 3DT3). Only key residues are shown for simplicity. Distance monitors (black dotted lines) show that the fluoro group is in close proximity to Met₄₂₁/Ile₃₇₃.

It is also instructive to discuss the selectivity of **23** and **24**. The bulky 1-naphthyl moiety most probably penetrates more deeply into the ER α Met₄₂₁/ ER β Ile₃₇₃ pocket and might be easier accommodated by the ER β Ile₃₇₃. However, the relatively low affinity of these compounds indicates that the distance between the aromatic 1-naphthyl group and the ER β Ile₃₇₃ may be too low to allow a favorable interaction and instead experience unfavorable interactions with both ER α Met₄₂₁ and ER β Ile₃₇₃.

Comparing the ER β selectivity of all designed THIQ analogues with respect to the binding affinity it can be noticed that electronegative and nonpolarizable N-substitutions favor a repulsive interaction with the ER α Met₄₂₁ side chain. For example, when compound **28** directs its electronegative atoms towards the electronegative S_δ atom of Met₄₂₁, electrostatic repulsion is unlikely to be offset by dispersive or inductive interactions, because it would be more difficult to further induce polarization. These effects would not be expected with the purely hydrophobic side chain of ER β Ile₃₇₃, making nonpolarizable functional groups like trifluoroacetamides or sulfonamides containing electronegative atoms a more attractive synthetic target than others with regard to improving ER β selectivity^[26d].

6.6 Conclusion

Based on the two-ring tetrahydroisoquinoline scaffold, for which no ER activity has been reported before^[11], a series of novel ER ligands have been successfully identified. The included nitrogen atom allowed for the rapid synthesis of simple analogues, including electronegative, charged and aromatic functional groups. The binding affinity toward both

ER α and ER β has been evaluated via fragment-based SAR and a variety of compounds featured significant agonistic binding affinities for both receptors. Incorporation of a trifluorosulfonamide group led to the identification of the best ER binder in this series, compound **28**, featuring an EC₅₀ of 0.59 μ M for the ER β and presenting a 4-fold higher selectivity toward ER β . A selected set of active compounds was further tested to determine their agonistic binding activity for the ERs in a cell-based transcriptional luciferase assay. Although the compounds featured a lower binding affinity for the ERs than the endogenous ligand estradiol, they actually act as good ligands for the ER activating the transcription of the target gene. Furthermore, the cell assay confirmed the observed ER β selectivity for compound **28**.

There has been much interest in finding potent and selective ligands for the ER β . Several recently described ER β selective ER agonists feature high binding affinities in the nanomolar region. These ligands are based on various scaffolds including biphenyls^[30a, 31], benzopyranes^[2, 26a, 32], naphthalenes^[30b, 33], tetrahydrochrysenes (THC)^[34], diarylpropionitriles (DPN)^[35], benzothiophenes^[36], isoxazoles^[37], benzoxazoles^[37b], benzoxazines^[38], benzofurans^[39], imidazoles/benzimidazoles^[37a], benzopyranes^[40], dihydrophenanthrenes^[41], fluorenones^[42], triazines^[43], indoles^[37a, 44], indenes/indenones^[45], quinolines^[46], steroidal^[47] and phytoestrogen analogous^[48]. Despite the around five magnitudes weaker binding affinity of the best binder in this series, **28**, this compound is a good hit. In particular, the affinity of compound **28** is notable with respect to the small size of the ligand. Most of the described compounds offer complex scaffolds with more than two rings. The structure of compound **28** was designed as simple as possibly with only two ring systems and the observed binding affinities for both ERs could be achieved without intensive optimization studies.

SAR and X-ray crystallographic studies demonstrated the importance of the presence and location of a phenolic 6-OH group within the THIQ scaffold that mimics the A-ring of estradiol and forms a hydrogen bond network that includes Glu₃₀₅ and Arg₃₄₆ of the ER β ligand binding cavity and a bound water molecule. The modest binding affinity of compound **28**, relatively to estradiol, can be explained due to the lack of a '17 β -OH group' mimic that is responsible for a further stabilizing hydrogen bond with His₄₇₅ and lowered hydrophobic potential, due to the smaller size of the amide group. An interesting follow up on the current class of molecules would therefore be a derivative of **28** that contains an additional OH-group and enables the interaction with His₄₇₅ and thus improves the affinity for ER β .

Further, it could be established that N-substitutions with electron-withdrawing groups enhanced ER affinity and were essential for ER β selectivity. These electronegative

substitutions provide access to ER α Met₄₂₁/ER β Ile₃₇₃, thus also playing a key role in the enhancement of ER β selectivity. The information from the SAR together with the established and optimized crystallization of the ER β LBD with variable THIQ analogues provides the entry to develop new compounds with maximized ER β affinity and selectivity.

6.7 Experimental Section

Chemical synthesis.

The synthesis of all described compounds was done by Willem van Otterlo (MPI Dortmund, Department of Chemical Biology and Molecular Sciences Institute, School of Chemistry, University of the Witwatersrand, Johannesburg).

1,2,3,4-tetrahydroisoquinoline (4) was purchased from SigmaAldrich.

1,2,3,4-Tetrahydro-6-isoquinolinol hydrobromide salt (10): Purchased from J&W Pharmed, Levittown, Philadelphia, USA.

Peptide synthesis. SRC-1 Box 2 peptide (⁶⁸⁵LTERHKILHRLQLQEGSPSD⁷⁰³) was synthesized from C- to N-terminus on solid support, using an automatic solid phase synthesizer (Syro II, Multisynth) on a 72 μ mol scale (150 mg of resin, loading of 0.7 mmol/g). The coupling of amino acids was carried out following the standard Fmoc-chemistry, using HOBt/DIC (4 equiv.) as amino acid activation, DMF as solvent and 4 equiv. of the protected Fmoc-amino acids. The resin was first swollen with DMF (1 x 20 min) and the Fmoc protecting group was removed by treatment with piperidine/DMF (2/3, 1 x 3 min; 1/4 1 x 10 min), then washed with DMF (6 x 1 min). One cycle of peptide elongation consisted of the following steps. First, the deprotected resin was treated for 50 min with a cocktail containing the appropriate amino acid (4 equiv, solution 0.3 M in DMF) with equimolar addition of HOBt and DIC (4 equiv, solution 0.3 M in DMF). After washing the resin with DMF (4 x 1 min), the Fmoc protecting group was removed by treatment with piperidine/DMF (2/3, 1 x 3 min; 1/4, 1 x 10 min). After deprotection, the resin was again washed with DMF (6 x 1 min). These steps were repeated until the peptide sequence was complete. After the completion of the sequence, the resin was subsequently washed with DMF (5 x 30 s), CH₂Cl₂ (5 x 30 s) and Et₂O (5 x 30 s) and dried under vacuum for 2-3 h. Cleavage and side chain deprotection was carried out by treatment of the resin for 3 h with a cleavage cocktail containing TFA/H₂O/EDT/TIS (94:2.5:2.5:1). The cleaved resin was washed with TFA (3 x 15 s) and the cleaved linear miniprotein was collected, concentrated by rotary co-evaporation with toluene into less than 1 mL solution and precipitated by addition of cold Et₂O (10 mL). The mixture was cooled in a liquid N₂ bath for 1 min, centrifuged (4000 rpm, 5 min, 4 °C) and the Et₂O was decanted from the pellet. Cold Et₂O was added again and the procedure was repeated twice. The crude peptide obtained was dissolved in H₂O/CH₃CN and lyophilized to dryness.

Peptide purification. The crude SRC-1 peptide was purified by reverse-phase HPLC on a Nucleodur C18 Gravity column (125 x 21 mm, Macherey-Nagel) with a linear gradient for 20 min. of A (0.1% TFA in H₂O) and B (0.1% TFA in CH₃CN) from 10% to 50% of B and flow rate of 25 mL min⁻¹ and were detected at 210 nm using a diode array UV/VIS detector. The identity and purity of the purified SRC-1 peptide was assessed by LC-

MS (ESI mass spectrometry). The expected peptide with a mass of was afforded with high purity (>85%) and 40 % yield. Following purification, the peptide was lyophilized and kept at $-20\text{ }^{\circ}\text{C}$.

Plasmids. The ER α LBD (residues 302-553) with an N-terminal His₆-tag was subcloned into pET15b and was a kind gift of A. Visser (Merck, before Organon). The construct pET15-hER β LBD expressing the ER β LBD (residues 260-502) with an N-terminal His₆-tag was constructed by subcloning an *NdeI/BamHI* fragment of pET15-hER β , provided by P. Donner (Bayer-Schering-Pharma AG), into pET15b (Novagen). Used primers: 5'-TTTTTTCATATGCTGGACGCCCTGAGCCCCGAGCAG-3' and 5'-TTTTTTGGATCCTCACCCGCGAAGCACGTGGGCATTCAGCATCTC-3'. The 0.24 kb fragment was digested with *NdeI/BamHI* and ligated with the newly double digested pET15b expression vector. The DNA fragment encoding for the ER β ₂₆₁₋₅₀₀ was amplified from pET15-hER β by PCR using the forward primer 5'-GAACCATGGACGACGCCCTGAGCCCCGAGCAGCTAGTG-3' and the reverse primer 5'-GGACTCGAGTTAGTCGTC AAGCACGTGGGCATTCAGCATCTC-3' to introduce the restriction site for *NcoI* and *XhoI*. The primers encode for three extra asp codons, one before the codon for D₂₆₁ and two after L₅₀₀. The expressed ER β LBD thus featured the following sequence: MD[D₂₆₁-L₅₀₀]DD. The PCR fragment was inserted into the *E. coli* expression vector pET16b (Novagen), double digested with the endonucleases *NcoI* and *XhoI*.

hER LBD Expression and Purification. The resulting plasmids (pET15-His-hER α LBD and pET15b-His-hER β LBD) were each transformed into *E. coli* BL21 (DE3) cells, and grown in 4 L TB medium using selection with Ampicillin (100 $\mu\text{g/ml}$). The cultures were incubated at $37\text{ }^{\circ}\text{C}$ till an OD₆₀₀ of ~ 1.2 and after cooling down to $15\text{ }^{\circ}\text{C}$, protein expression was induced by adding IPTG to a final concentration of 100 μM and the cells were grown in the presence of 10 μM E₂ (β -Estradiol, Serva) for 18-20 h at $18\text{ }^{\circ}\text{C}$ and harvested by centrifugation (4500 rpm, 20 min, $4\text{ }^{\circ}\text{C}$) and the pellet was stored at $-80\text{ }^{\circ}\text{C}$ till further use.

Cells were lysed with a micro fluidizer (4 passes at 600 kPa) in lysis buffer (PBS buffer containing 26.8 mM KCl, 14.7 mM KH₂PO₄, 78.1 mM Na₂HPO₄, 370 mM NaCl, 40 mM Imidazol, pH 8 and 10 % glycerol) and centrifuged (20.000 rpm, 30 min, $4\text{ }^{\circ}\text{C}$). The soluble cell lysate was immobilized on an equilibrated Nickel-NTA agarose column (HisTrap HP, 5 ml, Amersham Biosciences), washed with lysis buffer and eluted via a gradient using elution buffer (PBS buffer containing 26.8 mM KCl, 14.7 mM KH₂PO₄, 78.1 mM Na₂HPO₄, 370 mM NaCl, 500 mM Imidazol pH 8, 10% glycerol). Fractions containing the fusion protein were combined and desalted on a Sephadex G25 PD-10 column (Amersham Biosciences) using desalting buffer (20 mM Tris, 25 mM NaCl, 10 % glycerol and 0,05 % β -Octylglycosid). Determination of purification efficiency was performed by SDS gel electrophoresis. The concentration of both proteins was quantified using Nanodrop at a wavelength of 280 nm.

Expression and purification of the ER β LBD for crystallization. The ER β LBD₂₆₁₋₅₀₀ was overexpressed from high-density culture of *E. coli* BL21 (DE3) host cells (Statagene). Harvested cells were lysed using a micro fluidizer in a buffer containing 20 mM Tris-HCl pH 7.5, 0.5 M NaCl, 5 mM DTT, and 1 mM EDTA (10 ml/g of cells). Clarified lysate was flowed through a pre-equilibrated 3 ml estradiol-Sepharose column (PTI Research, Inc.) and washed with 250 ml of 10 mM Tris-HCl, pH 7.5 containing 0.5 M NaCl and 1 mM EDTA (buffer A). The column was then re-equilibrated with 50 ml of 10 mM Tris-HCl, pH 7.5, 0.2 M NaCl, and 1 mM EDTA (buffer B), and then the protein was carboxymethylated using 50 ml of buffer B containing 5 mM iodoacetic

acid. After incubation for 1 h at 4 °C, the column was washed by 400 ml of buffer A, followed by elution in 20 ml buffer A containing 50-200 μ M estradiol. Subsequently partial contaminations were removed by size exclusion chromatography (Sephadex 75, HiLoad 26/60, GE Healthcare Biosciences). Fractions containing the purified ER β LBD₂₆₁₋₅₀₀ were recombined and the buffer was exchanged into buffer B containing 5 mM DTT by passing the solution through a Sephadex G25 PD-10 column (Amersham Biosciences). Finally, the eluted ER β LBD₂₆₁₋₅₀₀ was concentrated to 10-12 mg/ml using Amicon ultra centrifuge tubes (MWCO 10 kDa) and characterized by SDS-PAGE and photometric determination of protein concentration using Nanodrop at a wavelength of 280 nm.

Estrogen Receptor Fluorescence Polarization Assay. A master solution in assay buffer (20 mM HEPES, pH 7.9; 500 mM NaCl; 1 mM EDTA; 10 % glycerol; 0.005% Nonidet P-40) was prepared containing ER α (20 μ M) or ER β (20 μ M) and 0.5 μ M fluorescently labelled peptide FI-CLTERHKILHRLLQEGSPSD (from SRC1 Box 2). A stock solution of the compounds was serially diluted (1.8-fold dilution) into assay buffer in a 384-well plate (Perkin Elmer, Optiplate-384 F) to a volume of 40 μ l in each well. To these wells 10 μ l of the master solution were added to obtain a final volume of 50 μ L with the following final concentrations: 4 μ M ER α or 4 μ M ER β and 0.1 μ M fluorescently labelled peptide FI-CLTERHKILHRLLQEGSPSD in addition to the compound concentration. The plate was centrifuged (5 min, 7000 rpm, 4 °C) and incubated for 1 h at 4 °C in the dark. The fluorescence polarization was measured at 23 °C using a plate reader (Safire-2, Tecan) with a 470 nm excitation and a 519 nm emission filter with 50 reads per well using a gain of 80 and a G-factor of 1.154. The normalized data were fitted with ORIGIN 7 (Scientific Graphing and Analysis Software, OriginLab Corp.) using nonlinear regression analysis with a sigmoidal dose-response equation to determine the K_D value of ligand binding. Pure buffer was used as the background signal. Wells with only peptide and buffer was used as control for minimum polarization. Measurements were done in triplicates.

Cell culture, cell transfection, and cell treatment. Human osteosarcoma U2OS cells were cultivated in complete Dulbecco's Modified Eagle Medium (DMEM) in the presence of 10% Fetal Bovine Serum (FBS) and Penicillin-Streptomycin (PS) at 37 °C under 5% CO₂. After reaching a confluence of 90%, the medium was aspirated and the cells were washed with PBS, trypsinized, and transferred to new DMEM. Before transfection, 80 μ l of the cells were transferred to 24-well plates in 500 μ l phenol red-free DMEM containing 5% Charcoal-treated FBS (DMEMdep; Hyclone) for ~16 h at 37 °C. For the Dual-Luciferase Reporter Assay System (Promega), cells were transfected with 10 ng hER α or hER β DNA, 200 ng ERE-LUC and 2 ng SV40 *Renilla* Luciferase as control in 50 μ l DMEMdep using 1 μ g/ μ l polyethylenimine (PEI; MW: 25 kDa; Polysciences). As a further control, cells were transfected with 700 ng of the pcDNA 3 empty vector to equalize the total amount of DNA per well. 8 h after transfection, cells were treated with fixed concentrations or serially dilutions of the desired ligand in DMEMdep including a maximum of 1% DMSO for about 16 h at 37°C before harvesting. Untreated cells were used as vehicle controls and incubation with estradiol was used as positive control. Cell lysis was realized by washing the cells with 1x PBS followed by an incubation step with 100 μ m lysis buffer (Promega) for 1 h with slight agitation.

Cell-based transcriptional assay. In the DLR assay, the activities of firefly and *Renilla* luciferase are measured sequentially from a single step. The firefly luciferase reporter is measured first by adding 100 μ l Luciferase

Assay Reagent II (LAR II; containing the substrate the firefly luciferase, beetle luciferin) to 20 μ l lysate to generate a stabilized luminescent signal. After quantifying the firefly luminescence using a microplate reader (Synergy HT, Bio-Tek) the reaction is quenched and the *Renilla* luciferase reaction is simultaneously initiated by dispensing 100 μ l of a Stop & Glo Reagent (containing the substrate for *Renilla*, coelenterazine) to the same wells. In order to minimize experimental variability the firefly activity was normalized against the *Renilla* luciferase luminescence.

$$N = \frac{PE(\textit{Firefly})}{PE(\textit{Renilla})} \quad (\text{Eq. 1})$$

where N is the normalized firefly luciferase activity and PE is the photon emission.

The percentage of the luciferase activity induced by the applied compound was determined using equation 2:

$$\% \textit{Activity} = \frac{(\textit{Firefly} \div \textit{Renilla})_{\textit{Compound}}}{(\textit{Firefly} \div \textit{Renilla})_{\textit{Control}}} \quad (\text{Eq. 2})$$

where %activity is the normalized firefly luciferase activity induced by the applied compound against the normalized firefly luciferase activity in the absence of ligand.

Crystallography. The ER β LBD - E₂ complex was concentrated to 12 mg/ml in a buffer containing 0.2 M NaCl, 1 mM EDTA, 5 mM DTT and 10 mM Tris-HCl at pH 7.5 and then mixed with the SRC-1 Box 2 peptide at a molecular ratio of 1.5 : 1 peptide to protein – ligand complex. The ER β LBD bound to compound **16** was concentrated to 16.13 mg/ml before incubating with the peptide. The compound **18** occupied ER β LBD was used at a final concentration of 19.27 mg/ml and the ER β LBD – compound **28** complex was concentrated to 12.43 mg/ml. Prior incubation with the ligand occupied ER β LBD, the peptide was dissolved in crystallization buffer to a concentration of 8.97 mM and extensively dialyzed against the same buffer for three days at 4 °C using a dialysis membrane from Spectra/Por (MWCO: 1 kDa). In order to crystallize the ER β LBD complex, initial screenings employing JCSG+, JCSG Core I, JCSG Core II, JCSG Core III and JCSG Core IV from Qiagen were performed at 20 and 4 °C using the sitting drop vapor diffusion. The protein concentration in the setups was 12 mg/ml; 0.1 μ l protein solutions were automatically mixed with 0.1 μ l reservoir solution in 96 well Corning pZero plates using a phoenix pipetting robot. The sitting drops were equilibrated against reservoirs with a volume of 70 μ l. After one to two days, hexagonal crystals appeared in several conditions typically containing 0.2 M of a diverse salt and 20% PEG 3350 and reaching a maximal size of 220 x 88 x 73 μ m within three days. Reproduction of these crystals were performed in 24 well EasyXtal DG-tools (Qiagen) using the hanging drop vapor diffusion method. Drops with a size of 2 to 5 μ l using a different reservoir to protein ratio were manually mixed at 4 °C and equilibrated against reservoirs with a volume of 1 ml. Optimal crystals were grown over night in 4.5 μ l drops with a protein solution to reservoir ratio of 3 : 1.5 in a condition containing 0.2 M Ammonium Acetate, 0.1 M BIS-TRIS pH 5.5 and 17% PEG10000 (v/v) at 4 °C. These crystals reached a final size of 0.5 x 0.4 x 0.3 mm within five days and were of fine quality as judged by light microscopy. After soaking in reservoir solution supplemented with glycerol to 20%, the crystals were cryo-cooled in liquid nitrogen for X-ray data collection at 100 K. Since the ER β LBD has been crystallized before and suitable protein structures could be

found in the Protein Data Bank, the crystal structure could be solved by molecular replacement of a ER β LBD/ligand/peptide complex (PDB code: 1u9e)^[26d] using the program PHASER^[49].

A native data set was collected from an optimized ER β LBD complex crystal at 100 K on a Nonius AXS MICRO star at a wavelength of 1.548 Å using a MAR dtb detector. The crystal-to-detector distance was 150 mm and the oscillation range was 1.0°. Data were indexed, integrated and scaled with the XDS package. The crystals diffracted to a maximal resolution of about 2.2 Å and a check for possible systematic absences revealed that the crystals belonged to space group P3₁, with one homo dimer per asymmetric unit. According to Matthews^[50], protein crystals typically exhibit a solvent content of 30 to 70%. Using the information about the volume of the unit cell and the molecular weight of the used protein complex, the number of molecules per asymmetric unit (ASU) can be estimated. The calculation of the Matthews parameters for a molecular weight of 29533 Da per ER β LBD/ligand/peptide complex and a cell volume of 475250 Å³, probably resulted in two complexes with a $V_M = 2.68 \text{ Å}^3/\text{Da}$ and a solvent content of 54.17% in an asymmetric unit. Data set statistics of all crystal structures are given in Table 3. The structure building was performed by calculation of the electron density maps using the program Coot^[51]. The iterative structure refinement was carried out with the program REFMAC^[52] from the program package CCP4 suite^[53]. All structural representations were prepared with pymol (www.pymol.org).

Table 3: X-ray data collection and statistics for the ER β LBD structures studies in this chapter

protein	ER β LBD	ER β LBD	ER β LBD	ER β LBD
ligand	E ₂	Compound 16	Compound 18	Compound 28
maximal resolution (Å)	2.2	2.3	2.4	2.2
space group	P3 ₁	P3 ₁	P3 ₁	P3 ₁
unit cell dimensions	a = 71.9 Å b = 71.9 Å c = 113.3 Å $\alpha = \beta = 90^\circ, \gamma = 120^\circ$	a = 71.9 Å b = 71.9 Å c = 113.3 Å $\alpha = \beta = 90^\circ, \gamma = 120^\circ$	a = 71.9 Å b = 71.9 Å c = 113.3 Å $\alpha = \beta = 90^\circ, \gamma = 120^\circ$	a = 71.9 Å b = 71.9 Å c = 113.3 Å $\alpha = \beta = 90^\circ, \gamma = 120^\circ$
number of molecules per asymmetric unit	2	2	2	2
wavelength (Å)	1.548	1.548	1.548	1.548
completeness (%)	98.6	99.7	98.3	98.9
R _{sym} (%)	27.9	31.3	19.4	21.1
mean I/ σ (I)	14.18	16.01	16.98	13.03

$R_{\text{sym}} = \sum |I(h)j - \langle I(h) \rangle| / \sum I(h)j$, where $I(h)j$ is the scaled observed intensity of the i^{th} symmetry related observation of the reflexes h and $\langle I(h) \rangle$ is the mean value.

6.8 References

- [1] M. Sato, T. A. Grese, J. A. Dodge, H. U. Bryant, C. H. Turner, *J Med Chem* **1999**, 42, 1.
- [2] G. G. Kuiper, B. Carlsson, K. Grandien, E. Enmark, J. Haggblad, S. Nilsson, J. A. Gustafsson, *Endocrinology* **1997**, 138, 863.
- [3] N. J. McKenna, R. B. Lanz, B. W. O'Malley, *Endocr Rev* **1999**, 20, 321.
- [4] a)V. C. Jordan, *J Med Chem* **2003**, 46, 1081; b)C. P. Miller, *Curr Pharm Des* **2002**, 8, 2089.
- [5] A. Haddow, J. M. Watkinson, E. Paterson, P. C. Koller, *British Medical Journal* **1944**, 1944, 393.
- [6] a)V. C. Jordan, *Cancer Cell* **2004**, 5, 207; b)J. A. Cauley, S. R. Cummings, D. M. Black, S. R. Mascioli, D. G. Seeley, *Am J Obstet Gynecol* **1990**, 163, 1438; c)D. C. Robinson, J. D. Bloss, M. A. Schiano, *Gynecologic Oncology* **1995**, 59, 186.
- [7] S. Green, P. Walter, V. Kumar, A. Krust, J. M. Bornert, P. Argos, P. Chambon, *Nature* **1986**, 320, 134.
- [8] a)G. L. Greene, P. Gilna, M. Waterfield, A. Baker, Y. Hort, J. Shine, *Science* **1986**, 231, 1150; b)G. G. Kuiper, E. Enmark, M. Pelto-Huikko, S. Nilsson, J. A. Gustafsson, *Proc Natl Acad Sci U S A* **1996**, 93, 5925; c)S. Mosselman, J. Polman, R. Dijkema, *Febs Letters* **1996**, 392, 49.
- [9] Z. Weihua, S. Andersson, G. Cheng, E. R. Simpson, M. Warner, J. A. Gustafsson, *FEBS Lett* **2003**, 546, 17.
- [10] a)S. Nilsson, G. Kuiper, J. A. Gustafsson, *Trends Endocrinol Metab* **1998**, 9, 387; b)O. B. Wallace, T. I. Richardson, J. A. Dodge, *Curr Top Med Chem* **2003**, 3, 1663; c)H. A. Harris, L. M. Albert, Y. Leathurby, M. S. Malamas, R. E. Mewshaw, C. P. Miller, Y. P. Kharode, J. Marzolf, B. S. Komm, R. C. Winneker, D. E. Frail, R. A. Henderson, Y. Zhu, J. C. Keith, Jr., *Endocrinology* **2003**, 144, 4241.
- [11] M. Manger, M. Scheck, H. Prinz, J. P. von Kries, T. Langer, K. Saxena, H. Schwalbe, A. Furstner, J. Rademann, H. Waldmann, *ChemBiochem* **2005**, 6, 1749.
- [12] S. Renner, W. A. van Otterlo, M. Dominguez Seoane, S. Mocklinghoff, B. Hofmann, S. Wetzel, A. Schuffenhauer, P. Ertl, T. I. Oprea, D. Steinhilber, L. Brunsveld, D. Rauh, H. Waldmann, *Nat Chem Biol* **2009**, 5, 585.
- [13] a)R. Paul, J. A. Coppola, E. Cohen, *Journal of Medicinal Chemistry* **1972**, 15, 720; b)K. Nagarajan, P. K. Talwalker, C. L. Kulkarni, R. K. Shah, S. J. Shenoy, S. S. Prabhu, *Indian Journal of Chemistry Section B-Organic Chemistry Including Medicinal Chemistry* **1985**, 24, 83; c)R. Chesworth, M. P. Zawistoski, B. A. Lefker, K. O. Cameron, R. F. Day, F. M. Mangano, R. L. Rosati, S. Colella, D. N. Petersen, A. Brault, B. Lu, L. C. Pan, P. Perry, O. Ng, T. A. Castleberry, T. A. Owen, T. A. Brown, D. D. Thompson, P. DaSilva-Jardine, *Bioorg Med Chem Lett* **2004**, 14, 2729; d)J. Renaud, S. F. Bischoff, T. Buhl, P. Floersheim, B. Fournier, M. Geiser, C. Halleux, J. Kallen, H. Keller, P. Ramage, *J Med Chem* **2005**, 48, 364; e)H. R. Lin, M. K. Safo, D. J. Abraham, *Bioorg Med Chem Lett* **2007**, 17, 2581.
- [14] R. L. Rosati, P. Da Silva Jardine, K. O. Cameron, D. D. Thompson, H. Z. Ke, S. M. Toler, T. A. Brown, L. C. Pan, C. F. Ebbinghaus, A. R. Reinhold, N. C. Elliott, B. N. Newhouse, C. M. Tjoa, P. M. Sweetnam, M. J. Cole, M. W. Arriola, J. W. Gauthier, D. T. Crawford, D. F. Nickerson, C. M. Pirie, H. Qi, H. A. Simmons, G. T. Tkalcovic, *J Med Chem* **1998**, 41, 2928.
- [15] P. AstraZeneca, **WO2002/46164**.
- [16] T. R. Hoye, M. K. Renner, T. J. VosDiNardo, *Journal of Organic Chemistry* **1997**, 62, 4168.
- [17] a)W. J. Checovich, R. E. Bolger, T. Burke, *Nature* **1995**, 375, 254; b)D. M. Jameson, W. H. Sawyer, *Biochemical Spectroscopy* **1995**, 246, 283; c)T. Heyduk, Y. Ma, H. Tang, R. H. Ebright, *Methods Enzymol* **1996**, 274, 492.
- [18] U. K. Laemmli, *Nature* **1970**, 227, 680.
- [19] R. B. Merrifield, *Journal of the American Chemical Society* **1963**, 85, 2149.
- [20] a)B. D. Darimont, R. L. Wagner, J. W. Apriletti, M. R. Stallcup, P. J. Kushner, J. D. Baxter, R. J. Fletterick, K. R. Yamamoto, *Genes Dev* **1998**, 12, 3343; b)S. E. Fawell, J. A. Lees, R. White, M. G. Parker, *Cell* **1990**, 60, 953.
- [21] J. Ponten, E. Saksela, *Int J Cancer* **1967**, 2, 434.
- [22] a)K. V. Wood, J. R. de Wet, N. Dewji, M. DeLuca, *Biochem Biophys Res Commun* **1984**, 124, 592; b)J. R. de Wet, K. V. Wood, D. R. Helinski, M. DeLuca, *Proc Natl Acad Sci U S A* **1985**, 82, 7870.
- [23] J. C. Matthews, K. Hori, M. J. Cormier, *Biochemistry* **1977**, 16, 85.
- [24] A. Kallio, T. Guo, E. Lamminen, J. Seppanen, L. Kangas, H. K. Vaananen, P. Harkonen, *Mol Cell Endocrinol* **2008**, 289, 38.
- [25] W. Kabsch, *Journal of Applied Crystallography* **1993**, 26, 795.
- [26] a)A. C. Pike, A. M. Brzozowski, R. E. Hubbard, T. Bonn, A. G. Thorsell, O. Engstrom, J. Ljunggren, J. A. Gustafsson, M. Carlquist, *EMBO J* **1999**, 18, 4608; b)A. K. Shiau, D. Barstad, P. M. Loria, L. Cheng, P. J. Kushner, D. A. Agard, G. L. Greene, *Cell* **1998**, 95, 927; c)A. M. Brzozowski, A. C. W. Pike, Z. Dauter, R. E. Hubbard, T. Bonn, O. Engstrom, L. Ohman, G. L. Greene, J. A. Gustafsson, M. Carlquist, *Nature* **1997**, 389, 753; d)E. S. Manas, R. J. Unwalla, Z. B. Xu, M. S. Malamas, C. P. Miller, H. A. Harris, C. Hsiao, T. Akopian, W. T. Hum, K. Malakian, S. Wolfrom, A. Bapat, R. A. Bhat, M. L. Stahl, W. S. Somers, J. C. Alvarez, *J Am Chem Soc* **2004**, 126, 15106.

- [27] a)A. C. Pike, A. M. Brzozowski, R. E. Hubbard, *J Steroid Biochem Mol Biol* **2000**, 74, 261; b)A. Warnmark, E. Treuter, J. A. Gustafsson, R. E. Hubbard, A. M. Brzozowski, A. C. W. Pike, *Journal of Biological Chemistry* **2002**, 277, 21862.
- [28] G. M. Anstead, K. E. Carlson, J. A. Katzenellenbogen, *Steroids* **1997**, 62, 268.
- [29] M. S. Malamas, E. S. Manas, R. E. McDevitt, I. Gunawan, Z. B. Xu, M. D. Collini, C. P. Miller, T. Dinh, R. A. Henderson, J. C. Keith, H. A. Harris, *Journal of Medicinal Chemistry* **2004**, 47, 5021.
- [30] a)C. Yang, R. Edsall, Jr., H. A. Harris, X. Zhang, E. S. Manas, R. E. Mewshaw, *Bioorg Med Chem* **2004**, 12, 2553; b)R. E. Mewshaw, R. J. Edsall, Jr., C. Yang, E. S. Manas, Z. B. Xu, R. A. Henderson, J. C. Keith, Jr., H. A. Harris, *J Med Chem* **2005**, 48, 3953.
- [31] R. J. Edsall, Jr., H. A. Harris, E. S. Manas, R. E. Mewshaw, *Bioorg Med Chem* **2003**, 11, 3457.
- [32] a)T. Barkhem, B. Carlsson, Y. Nilsson, E. Enmark, J. Gustafsson, S. Nilsson, *Mol Pharmacol* **1998**, 54, 105; b)F. Roelens, N. Heldring, W. Dhooge, M. Bengtsson, F. Comhaire, J. A. Gustafsson, E. Treuter, D. De Keukeleire, *J Med Chem* **2006**, 49, 7357; c)M. De Angelis, F. Stossi, M. Waibel, B. S. Katzenellenbogen, J. A. Katzenellenbogen, *Bioorg Med Chem* **2005**, 13, 6529.
- [33] R. E. Mewshaw, S. M. Bowen, H. A. Harris, Z. B. Xu, E. S. Manas, S. T. Cohn, *Bioorg Med Chem Lett* **2007**, 17, 902.
- [34] a)M. J. Meyers, J. Sun, K. E. Carlson, B. S. Katzenellenbogen, J. A. Katzenellenbogen, *J Med Chem* **1999**, 42, 2456; b)J. Sun, M. J. Meyers, B. E. Fink, R. Rajendran, J. A. Katzenellenbogen, B. S. Katzenellenbogen, *Endocrinology* **1999**, 140, 800; c)A. K. Shiau, D. Barstad, J. T. Radek, M. J. Meyers, K. W. Nettles, B. S. Katzenellenbogen, J. A. Katzenellenbogen, D. A. Agard, G. L. Greene, *Nat Struct Biol* **2002**, 9, 359.
- [35] a)M. J. Meyers, J. Sun, K. E. Carlson, G. A. Marriner, B. S. Katzenellenbogen, J. A. Katzenellenbogen, *J Med Chem* **2001**, 44, 4230; b)J. Sun, J. Baudry, J. A. Katzenellenbogen, B. S. Katzenellenbogen, *Mol Endocrinol* **2003**, 17, 247.
- [36] U. Schopfer, P. Schoeffter, S. F. Bischoff, J. Nozulak, D. Feuerbach, P. Floersheim, *J Med Chem* **2002**, 45, 1399.
- [37] a)R. Chesworth, M. D. Wessel, L. Heyden, F. M. Mangano, M. Zawistoski, L. Gegnas, D. Galluzzo, B. Lefker, K. O. Cameron, J. Tickner, B. Lu, T. A. Castleberry, D. N. Petersen, A. Brault, P. Perry, O. Ng, T. A. Owen, L. Pan, H. Z. Ke, T. A. Brown, D. D. Thompson, P. DaSilva-Jardine, *Bioorg Med Chem Lett* **2005**, 15, 5562; b)M. S. Malamas, E. S. Manas, R. E. McDevitt, I. Gunawan, Z. B. Xu, M. D. Collini, C. P. Miller, T. Dinh, R. A. Henderson, J. C. Keith, Jr., H. A. Harris, *J Med Chem* **2004**, 47, 5021.
- [38] W. Yang, Y. Wang, Z. Ma, R. Golla, T. Stouch, R. Seethala, S. Johnson, R. Zhou, T. Gungor, J. H. Feyen, J. K. Dickson, Jr., *Bioorg Med Chem Lett* **2004**, 14, 2327.
- [39] M. D. Collini, D. H. Kaufman, E. S. Manas, H. A. Harris, R. A. Henderson, Z. B. Xu, R. J. Unwalla, C. P. Miller, *Bioorg Med Chem Lett* **2004**, 14, 4925.
- [40] a)B. H. Norman, J. A. Dodge, T. I. Richardson, P. S. Borromeo, C. W. Lugar, S. A. Jones, K. Chen, Y. Wang, G. L. Durst, R. J. Barr, C. Montrose-Rafizadeh, H. E. Osborne, R. M. Amos, S. Guo, A. Boodhoo, V. Krishnan, *J Med Chem* **2006**, 49, 6155; b)T. I. Richardson, B. H. Norman, C. W. Lugar, S. A. Jones, Y. Wang, J. D. Durbin, V. Krishnan, J. A. Dodge, *Bioorg Med Chem Lett* **2007**, 17, 3570.
- [41] a)W. Sun, L. D. Cama, E. T. Birzin, S. Warriar, L. Locco, R. Mosley, M. L. Hammond, S. P. Rohrer, *Bioorg Med Chem Lett* **2006**, 16, 1468; b)J. Pandey, A. K. Jha, K. Hajela, *Bioorg Med Chem* **2004**, 12, 2239.
- [42] a)R. R. Wilkening, R. W. Ratcliffe, E. C. Tynebor, K. J. Wildonger, A. K. Fried, M. L. Hammond, R. T. Mosley, P. M. Fitzgerald, N. Sharma, B. M. McKeever, S. Nilsson, M. Carlquist, A. Thorsell, L. Locco, R. Katz, K. Frisch, E. T. Birzin, H. A. Wilkinson, S. Mitra, S. Cai, E. C. Hayes, J. M. Schaeffer, S. P. Rohrer, *Bioorg Med Chem Lett* **2006**, 16, 3489; b)J. L. Wright, T. F. Gregory, S. R. Kesten, P. A. Boxer, K. A. Serpa, L. T. Meltzer, L. D. Wise, S. A. Espitia, C. S. Konkoy, E. R. Whittemore, R. M. Woodward, *J Med Chem* **2000**, 43, 3408; c)R. R. Wilkening, R. W. Ratcliffe, A. K. Fried, D. Meng, W. Sun, L. Colwell, S. Lambert, M. Greenlee, S. Nilsson, A. Thorsell, M. Mojena, C. Tudela, K. Frisch, W. Chan, E. T. Birzin, S. P. Rohrer, M. L. Hammond, *Bioorg Med Chem Lett* **2006**, 16, 3896; d)D. L. Parker, Jr., D. Meng, R. W. Ratcliffe, R. R. Wilkening, D. M. Sperbeck, M. L. Greenlee, L. F. Colwell, S. Lambert, E. T. Birzin, K. Frisch, S. P. Rohrer, S. Nilsson, A. G. Thorsell, M. L. Hammond, *Bioorg Med Chem Lett* **2006**, 16, 4652.
- [43] B. R. Henke, T. G. Consler, N. Go, R. L. Hale, D. R. Hohman, S. A. Jones, A. T. Lu, L. B. Moore, J. T. Moore, L. A. Orband-Miller, R. G. Robinett, J. Shearin, P. K. Spearing, E. L. Stewart, P. S. Turnbull, S. L. Weaver, S. P. Williams, G. B. Wisely, M. H. Lambert, *J Med Chem* **2002**, 45, 5492.
- [44] J. W. Ullrich, R. J. Unwalla, R. R. Singhaus, Jr., H. A. Harris, R. E. Mewshaw, *Bioorg Med Chem Lett* **2007**, 17, 118.

- [45] a)S. O. Mueller, J. M. Hall, D. L. Swope, L. C. Pedersen, K. S. Korach, *J Biol Chem* **2003**, 278, 12255; b)R. E. McDevitt, M. S. Malamas, E. S. Manas, R. J. Unwalla, Z. B. Xu, C. P. Miller, H. A. Harris, *Bioorg Med Chem Lett* **2005**, 15, 3137.
- [46] a)W. Chen, Z. Lin, M. Ning, C. Yang, X. Yan, Y. Xie, X. Shen, M. W. Wang, *Bioorg Med Chem* **2007**, 15, 5828; b)T. Gungor, Y. Chen, R. Golla, Z. Ma, J. R. Corte, J. P. Northrop, B. Bin, J. K. Dickson, T. Stouch, R. Zhou, S. E. Johnson, R. Seethala, J. H. Feyen, *J Med Chem* **2006**, 49, 2440; c)A. T. Vu, A. N. Campbell, H. A. Harris, R. J. Unwalla, E. S. Manas, R. E. Mewshaw, *Bioorg Med Chem Lett* **2007**, 17, 4053.
- [47] a)C. Hegele-Hartung, P. Siebel, O. Peters, D. Kosemund, G. Muller, A. Hillisch, A. Walter, J. Kraetzschmar, K. H. Fritzeimer, *Proc Natl Acad Sci U S A* **2004**, 101, 5129; b)Z. Weihua, R. Lathe, M. Warner, J. A. Gustafsson, *Proc Natl Acad Sci U S A* **2002**, 99, 13589; c)T. R. Pak, W. C. Chung, T. D. Lund, L. R. Hinds, C. M. Clay, R. J. Handa, *Endocrinology* **2005**, 146, 147; d)T. A. Blizzard, C. Gude, J. D. Morgan, 2nd, W. Chan, E. T. Birzin, M. Mojena, C. Tudela, F. Chen, K. Knecht, Q. Su, B. Kraker, R. T. Mosley, M. A. Holmes, N. Sharma, P. M. Fitzgerald, S. P. Rohrer, M. L. Hammond, *Bioorg Med Chem Lett* **2006**, 16, 834; e)T. A. Blizzard, C. Gude, W. Chan, E. T. Birzin, M. Mojena, C. Tudela, F. Chen, K. Knecht, Q. Su, B. Kraker, M. A. Holmes, S. P. Rohrer, M. L. Hammond, *Bioorg Med Chem Lett* **2007**, 17, 2944.
- [48] C. P. Miller, M. D. Collini, H. A. Harris, *Bioorg Med Chem Lett* **2003**, 13, 2399.
- [49] A. J. McCoy, R. W. Grosse-Kunstleve, L. C. Storoni, R. J. Read, *Acta Crystallogr D Biol Crystallogr* **2005**, 61, 458.
- [50] B. W. Matthews, *J Mol Biol* **1968**, 33, 491.
- [51] P. Emsley, K. Cowtan, *Acta Crystallogr D Biol Crystallogr* **2004**, 60, 2126.
- [52] G. N. Murshudov, A. A. Vagin, E. J. Dodson, *Acta Crystallogr D Biol Crystallogr* **1997**, 53, 240.
- [53] S. Bailey, *Acta Crystallographica Section D-Biological Crystallography* **1994**, 50, 760.

Modulating the Nuclear Receptor – Cofactor Interaction: Characterization and Inhibition

The nuclear receptor (NR) – coactivator interaction is one of the key steps in transcription control and concomitantly related to many diseases and an attractive drug target. Various ligands have been synthesized in the past to indirectly modulate this protein-protein interaction. However, despite their therapeutically value, the occurrence of undesired side effects entailed the requirement of novel approaches to address this issue. Direct inhibition of the NR – coactivator interaction by peptides and small molecules that compete with the coactivators for the NR binding groove represents an attractive new approach to increase the physiological knowledge on the complex processes involved in selective cofactor binding and might ultimately pave ways for novel drug development. Additionally, the control mechanisms beyond the ligand binding event, such as post-translational modifications, have come forward as important regulators of the NR – coactivator interaction for which molecular insights are urgently required.

Targeting the Nuclear Receptor – Coactivator interaction

Potent and selective peptide-based coactivator binding inhibitors (CBIs) have been developed in the past. The development of an on-bead peptide library screen for the identification of novel peptide inhibitor sequences for the AR - coactivator interaction newly described here allowed for the screening of peptides with non-proteinogenic amino acids in linear peptide sequences. Beads amendable to organic synthetic modifications and compatible to protein screening conditions were modified in a combinatorial fashion with a specific peptide library, leading to an One-Bead-One-Compound library. The library was incubated with Texas Red or Qdot labeled AR LBD and bright beads observed under the confocal fluorescence microscope were isolated to analyze their bound peptide sequence. The reliability of the assay as well as the exact binding affinity of the re-synthesized peptide hits for the AR LBD was investigated via fluorescence polarization studies (Figure 1). This On-Bead library screening method generated novel peptide sequences in a rapid manner, including non-natural amino acids that are able to inhibit the AR-cofactor interaction. Further, this methodology allows a screening of complete randomized libraries and other nuclear receptors such as the ER. This will hopefully result in a set of molecular tools that can be used in follow-up studies to help understanding critical issues such as selectivity and recognition motifs.

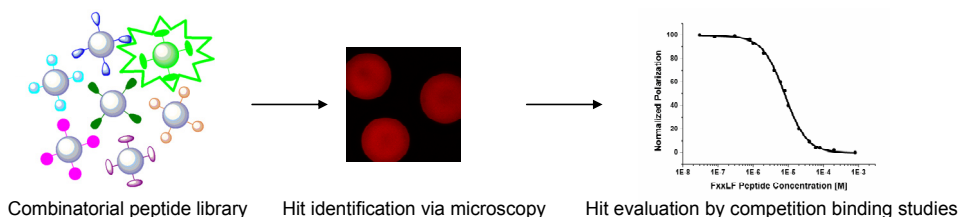


Figure 1: General method for On-Bead screening. Synthesis of an On-Bead peptide library using the AR FXXLF motif as scaffold, incubation with fluorescent AR LBD and detection of hits under the microscope, followed by isolation and identification of the peptide sequence. The reliability and the binding affinity were analyzed by fluorescence polarization competition studies.

A structural analysis of the AR-cofactor interaction shows that the coactivator peptide has to fold into a short helix to bind between a charge clamp at a fixed position on the AR surface. The length and stability of the helix thus appear to be crucial elements to achieve optimal binding. Control over helix stability and length is typically difficult for regular peptides and therefore stabilization of short peptides in a defined and stable fold is being pursued. Several different miniproteins featuring stabilized helices by means of two or three disulfide bridges were taken as scaffold for the *in silico* design of helical peptide binders for AR (Figure 2). Competitive fluorescence polarization binding studies for the AR LBD showed that many of the designed and synthesized miniproteins featured remarkable binding affinities around and below 1 μM . The length of the miniprotein helix was shown to be optimal when featuring around two turns. Since the influence of the point mutations on helix stability is significantly less prominent in miniproteins as in linear peptides, the peptides also allowed the evaluation of specific point mutations on binding affinity. Cysteine to methionine mutations of the miniproteins demonstrated the importance of the FXXLF motif on a performed stable helical segment in the miniprotein for high affinity binding. The introduction of an LXXLL motif into the helix enabled one miniprotein to bind the ER LBD. Initial crystallization studies of the miniproteins in complex with NR LBDs provide an entry to evaluate the molecular interactions of the miniproteins with the NR. Future attempts to generate miniproteins, stable under cellular conditions offer the possibility for applications in cell-based studies. As such, other miniprotein libraries can give even more detailed molecular insights into the molecular recognition of NRs by coactivators and provide the molecular requirements to generate new coactivator binding inhibitors for NRs.

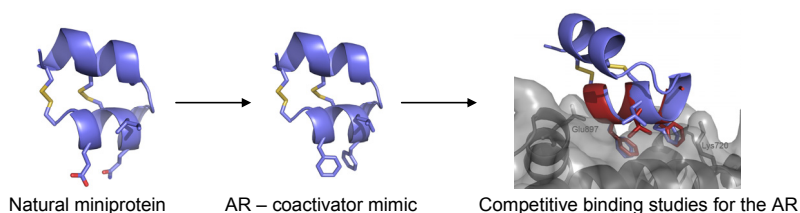


Figure 2: Strategy to generate AR binders by insertion of the FXXLF motif in the helix of a natural miniprotein. The *in silico* AR binding model was evaluated by competitive FP binding studies.

Chemical Biology Evaluations of Estrogen Receptor Post-translational Modifications

NRs undergo a variety of post-translational modifications (PTM). The influence of these PTMs on ligand binding and on the formation of multiple protein complexes is, however, largely unknown. A synthetic entry into NR constructs featuring specifically introduced PTMs, would open up the possibility to study the crosstalk between NRs and their cofactors, as well as other phenomena, on the molecular level. A protein semi-synthesis method to generate correctly folded and active ER α and ER β LBDs was successfully established. This method allowed the generation of the ER LBD with a selectively and homogeneously phosphorylated tyrosine, not accessible via the typical biochemical enzymatic approach. Using expressed protein ligation (EPL), the recombinant ER thioesters, lacking helix 12, were generated. Helix 12 of both isoforms was chemically synthesized bearing a phosphorylated tyrosine. The chosen strategy allows the use of a native cysteine in the ER LBD for successful ligation of the synthetic peptide to the protein thioester. Using this approach, both the phosphorylated ER α (pY₅₃₇) and the phosphorylated ER β LBD (pY₄₈₈) could be successfully synthesized.

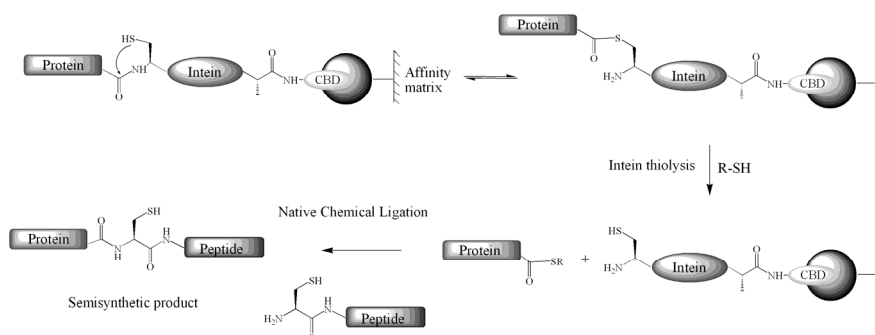


Figure 3: Purification strategy of the protein corpus as α -thioester and solid phase synthesis of the phosphorylated peptide. Subsequent ligation of both parts generates the phosphorylated ER LBD.

For the first time, the defined three-dimensional crystal structure of a phosphorylated ER β LBD in complex with cofactor peptide and the agonist estradiol (E₂) could be determined with a maximal resolution of 1.5 Å. The observation that the phosphorylated Y₄₈₈ sticks out of the global protein complex supports the assumption that the phosphorylated tyrosine can be target by the SH2 domain of cSrc via this phosphate. Further the access to crystal structures of post-translational modified NRs provides an entry to study the structural influence of PTMs on NR – protein interactions.

The influence of tyrosine phosphorylation on cofactor binding was investigated using an on-chip- and FRET-based cofactor recruitment studies. Tyrosine phosphorylation is able to decrease the binding efficiency of distinct coactivator peptides to agonist-bound ER α and ER β . This observation strongly suggests that tyrosine phosphorylation of both ER LBDs represents an important control site that is involved in regulating cofactor binding under certain cellular conditions. However, in contrast to ER α , tyrosine phosphorylation of the ER β appears to lead to a higher flexible character of helix 12. This is accompanied by a significant increased affinity to coactivator peptides in the absence of ligand or in the presence of antagonist, as most prominent effect. Interestingly, this agonist-independent increase in cofactor binding was not observed for the ER α , reinforcing the idea that phosphorylation provides the two ER subtypes with distinct cofactor regulatory functions.

Fragment-based Design and Structural Elucidation of Estrogen Receptor Agonists

Despite the recent advances in directly inhibiting the NR – coactivator interaction with CBIs, the development of novel synthetic hormone analogues, binding the ligand binding pocket of the NR LBD, is of high importance, e.g. to generate sub-type selective ligands. Although many biological active compounds have been identified to bind the ER LBD, the search for compounds that selectively distinguish between the two ER isoforms ER α and ER β is still a challenging issue. Based on a simple two ring tetrahydroisoquinoline scaffold, a series of novel active scaffolds were identified. Using fragment based structure optimization it was demonstrate that certain functional groups incorporated during the optimization were important to enhance ER activity in both, biochemical and cellular transactivation assays up to an EC₅₀ value of 0.59 μ M and a 4-fold higher selectivity for ER β . Co-crystallization studies of some active compounds in complex with the ER β LBD revealed that the tetrahydroisoquinoline scaffold mimics the A and B ring of E₂ and thus forms hydrogen bonds with Glu₃₀₅/Arg₃₄₆ (Figure 4). Further, electronegative groups like trifluorosulfonamide groups enhance the affinity to ER, most likely via electrostatic interactions. Further, the observed

slight preference for ER β could be explained by electron-withdrawing groups, like fluoro groups, in close proximity to the Ile₃₇₃ residue. It seems that in ER α these groups can be less effectively accommodated due to the presence of the more bulky Met₄₂₁ at the same position. The identified new ER ligands together with the information gained from the x-ray studies present a suitable tool for further investigations of ER modulators. Especially with respect to the Ile₃₇₃/Met₄₂₁ interaction, these scaffolds could pave the way for the design of highly selective ER β binders.

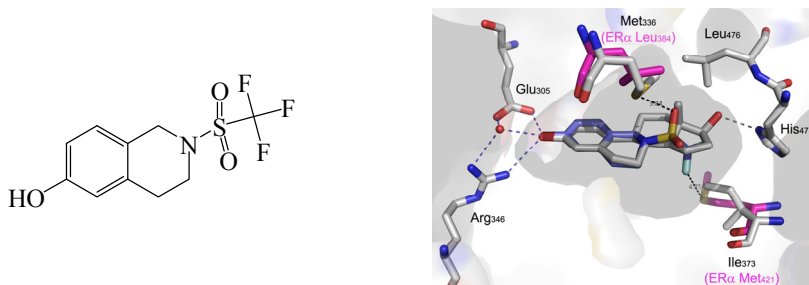


Figure 4: Chemical and co-crystal structure of the best found agonist in this study (blue) complexed with the ER β LBD and overlaid with E₂ (white). Residues of ER α that differ from the ER β ligand binding pocket were shown in pink (PBD 3DT3). Only key residues of the binding pocket are shown for simplicity.

Curriculum Vitae



Sabine Möcklinghoff was born on March 23th, 1981 in Dortmund, Germany. She attended her secondary education at the Stadtgymnasium Dortmund where she obtained the ‘Abitur’ (pre-university education) in 2000. In the same year she started studying Biology at the Ruhr-Universität Bochum. This study was completed with a graduation project (Diploma thesis) on molecular and physiological investigations on plastidic plant lipases in the laboratory of Plant Physiology under supervision of Prof. Dr. E.W. Weiler in 2005. In February 2006, she started as a PhD candidate at the Chemical Genomics Centre of the Max-Planck Society in Dortmund under supervision of Prof. Dr. Luc Brunsveld and finished the research at the Technische Universiteit Eindhoven in 2010. The aim of her research was to modulate the Nuclear Receptor – Cofactor Interaction using Chemical Biology approaches. The most important results of this investigation are presented in this thesis.

Acknowledgement

During the last four years I worked together with a lot of different people that contributed a great deal to this thesis and I would like to take the advantage to acknowledge them here.

First of all, I sincerely thank my supervisor prof.dr. Luc Brunsveld. Luc, I'm glad that you gave me the opportunity to work in your group as the first biologist. Your interest in and enthusiasm for the nuclear receptor project, the amicable working relationship, the many activities beyond the lab, the chance to learn so many different facets of the chemical biology field, your support and your confidence in my work always inspired me and made sure that I never lost the joy in my project and science in general. The past years were a fantastic experience that I don't want to miss and that will always remain in special memory.

I'm grateful for the referees of this thesis prof.dr. E. W. Meijer, prof.dr. H. Vader, dr. E. Kalkhoven and Prof. Dr. H. Mootz for their time and interest in my work and their constructive advices. Furthermore, I would like to thank dr. M. Merkx, prof.dr. K. Nicolay and prof.dr. C.A.A. van Boeckel for complementing the opposition during my defence.

A special gratitude is directed towards all co-workers that contributed to the success of this thesis. I want to especially highlight Dr. Belén Vaz Araujo. It was great fun working with you and setting up the lab from scratch together with you and Dr. Maarten Broeren. I thank Belén and Dr. Simon Folkertsma for the *in silico* design and the synthesis of the miniproteins. Further, I thank Belén, Dr. Dorothee Jonkheijm and Dr. Hülya Göksel for the synthesis of the on-bead peptide libraries. Dr. Maëlle Carraz and Wencke Adriaens are heartily thanked for their help with the synthesis of the peptides for the chemical ligation.

Arie Visser is gratefully acknowledged for giving me the possibility to spend a few weeks in his lab at Organon in Oss. Karin Smits and Elisabeth van der Vaart are thanked for their help with the FRET assay and the enzymatic ER phosphorylation assay. I also would like to thank Joost Uitdehaag for the possibility to perform my first protein crystallization trials in his lab.

Ich danke Prof. Dr. P. Donner, Dr. V. Pütter and Dr. H. Weinmann von Bayer-Schering Pharma für ihr Interesse an meiner Arbeit, anregende Diskussionen, Ratschläge und ihre Hilfe während meines kurzen Aufenthaltes in Berlin.

Mein herzlicher Dank gilt Dr. Christian Ottmann, Dr. Rolf Rose und Dr. Michael Weyand für die große Hilfe bei der Proteinkristallisierung. Es war eine Freude mit euch zu Arbeiten und von euch zu Lernen.

I'm grateful to Prof. Dr. Willem van Otterlo for the enjoyable collaboration, the synthesis of the THIQ analogues and his constant enthusiasm and interest in the project.

My gratitude goes to the whole work group at the technical university in Eindhoven for their kindness and for making me feel welcome each time I came to the Netherlands. Besonders danke ich Sascha Fuchs, meinem ehemaligen Masterstudenten und zukünftigen Paranimfen für die gute Zusammenarbeit, die Hilfe bei dem Zellassay und viele fröhliche Stunden. Weiterhin danke ich Inga Tharun für die tolle Zusammenarbeit und die phantastische Zeit auf der Konferenz in Cavtat. Gerne gebe ich das PTM Projekt in deine Hände weiter.

I was fortune to have had excellent lab-mates in Dr. Belén Vaz, Dr. Marta Domínguez, Dr. Jacqui Young, Dr. Hülya Göksel, Dr. Maarten Broeren, Dr. Hoang Nguyen, Dr. Trang Phan, Katja Petkau, Dr. Dorothee Jonkheijm, Dr. Maëlle Carraz and Tuyen Nguyen. I'm grateful for the kind teamwork and the amicable working atmosphere. I enjoyed working and spending countless funny leisure time events with you.

My fantastic room mates Sabine Klüter, Beate Aust, Jérôme Clerc, Sabine Wulfert, Zhe Ming Wang and Patrick Hauske are heartily thanked for the wonderful and relaxed atmosphere in the office and many awesome moments. Besonders danke ich Sabine Klüter für ihre Freundschaft und die tolle Zeit in München und Ringberg. Deine emotionale Unterstützung, sowie unsere regelmäßigen Tratschstunden haben für die nötige Zerstreuung und Regeneration gesorgt. Mein herzlicher Dank gilt auch Patrick Hauske für seine unendliche Geduld und selbstlose Hilfe bei der Peptidchemie und für die vielfältigen Unterhaltungen und Erheiterungen während so mancher Nachtschicht im CGC. Ohne deine Freundschaft wäre mir die Doktorandenzeit deutlich schwerer gefallen. Bleib wie du bist!

All remaining co-workers of the CGC and the MPI in Dortmund are gratefully acknowledged for the manifold experimental help and the various scientific and practical advices. Beside the lab work there was always time for recreation and fun in terms of countless cheerful coffee and cake breaks, common conferences, extensive parties, fantastic trips and sportive activities. I will miss you all!

Zum Schluß möchte ich noch meiner Mama und Philipp dafür danken, daß sie mich in Allem so liebevoll unterstützt und immer an mich geglaubt haben. Ich bin froh und stolz, daß ihr meine Familie seid. Sascha, ich danke dir für die bedingungslose Unterstützung während des gesamten Studiums und daß du auch in schweren Zeiten immer zu mir gehalten hast. Ohne deine Liebe und Geduld hätte ich es nicht geschafft. Du bist Alles für mich!

Sabine