Title
Discover coactivator proteomimetic inhibitors

Author
Geistlinger, Timothy R

Publication Date
2004

Peer reviewed|Thesis/dissertation

Discovering Coactivator Proteomimetic Inhibitors

by

Timothy R. Geistlinger

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Chemistry and Chemical Biology

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Date University Librarian

Degree Conferred:.....

Dedication

This work is dedicated to friends and family. You have all been an inspiration and a great deal of support. If everyone were only so lucky, maybe our times would be filled with less fear and more confident benevolence.

Acknowledgements

The completion of this research project is just as much a relief, as it is a privilege. Tantamount to the research itself, this work represents a significant period of time in my life when I had the opportunity to work with the most amazing collection of scientists and friends any person could ever hope for. It is due to many, if not all of you, that I have the opportunity to tell this story, and the honor of graduating from UCSF with a doctorate degree. You have all played a bigger role in this than you know and I promise you that I will do all that I can to be worthy.

This research project began over five years ago when it was stenciled on the back of an envelope and revealed to me by a few professors who were willing to give me a chance. They encouraged me to explore the possibility of doing what so many had tried and failed or said was impossible, and to do it differently. Robert Fletterick pulled me into his office one day to talk about my life, skiing, and science,.. a conversation for which he is quite famous. At some magical point, Robert asked me if I wanted to work with one of the greatest chemists in the world to explore an idea that was very similar to a proposal I had recently written. He was referring to Kip Guy and the project that would consume the next five years of my life. Thank you Flett, you were right about Kip. And thank you for your unbiased professional advice over the years. Shortly thereafter, I walked into a newly carpeted lab. It had a lot of empty shelves, one computer, and Kip Guy. After we met and discussed several potential projects in the lab, Kip laughed in his very Kip sort of way and said, "Oh yeah,.. You're going to have to buy everything before you can get started." I was Kip's first student and Kip was my first PhD mentor. Kip turned out to be quite possibly the nicest, most eclectic, brilliant, well-rounded, sane professor I had ever met, and one who just happened to be only two years older than me. Thank you for sharing your patience, energy and wisdom with me Kip. You have taught me a hell of a lot. Holly Ingraham has been a continual educator, supporter, and a very good influence. She leads by asking and successfully exploring novel questions to biological networks in any way that she can. She consistently manages to make me rethink paradigms I thought I already understood. And she offers excellent advice, always with a smile, and even in the wee hours of a Gordon Conference while simultaneously

taking on the likes of David Moore or Jeff Rosen. Keith Yamamoto is the man who scared the hell out of me and forced me to ask myself if I was in right place. I will never forget that and I thank you for all of your support over the years. Peter Kollman, in life he advised me on the dynamic potentials in my molecules and in memory he made me a better, more dynamic, human being. I miss his booming voice and our walks home together down Parnassus Avenue. I am grateful to the Friday morning ER meetings with Peter Kushner and the excellent collection of amazing scientists who always brought me closer to the cellular biology and the clinically relevant questions of my research. All of these professors are pioneers in their fields, addressing some of the most difficult questions in medical science, and each of them took valuable time out of their schedules to work with me. Thank you!

Three scientists with whom I have had the great pleasure of knowing as friends and colleagues also made several of the critical, seminal discoveries formulating the foundation for my research hypothesis. Beatrice Darimont, undoubtedly one of the greatest molecular biologists I have ever met, made the biochemical discoveries that this interface could be reduced to the small NR box sequences, and helped me with conducting several of my first experiments. Together with Richard Wagner, they elucidated the structural clues behind this interface with the thyroid hormone receptor. Andrew Shiau, positively one of the premiere structural biologists in our times, solved the first structures of the estrogen receptor showing the agonist and antagonist conformations. And now more recently, without Andy's help in my latest efforts to solve the crystal structures of these inhibitors with the estrogen receptor I would not have been able to see what my inhibitors actually looked like.

No UCSF graduate career would be possible without the willingness of the UCSF community to collaborate, share space and equipment, and cross-pollinate scientific ideas. During my studies I spent a considerable amount of time in each of the following laboratories for which I want to say thank you to everyone for putting up with me. In no particular order I thank the labs of: the Irwin Kuntz for computational time and know-how; David Agard for crystallography, circular dichroism and protein purification; the Shigeaki Kato lab at the University of Tokyo for allowing me to conduct rather unconventional cellular assays and letting me give them all a hug each morning; the

Keith Yamamoto lab for biochemistry tools and knowledge; Charles Craik for discussion and allowing me to steel piperidine in the middle of the night; John Baxter, Fred Schaufele and Paul Web who patiently put up with so many of my questions and ideas; Volker Doetsch for NMR studies; Matt Bogyo's for not loosing it when I fried the detector on their brand new LC; the CAT lab; and, of course everyone in the Guy lab who put up with my perpetual messes, drawn-out presentations, my not so funny sense of humor, and managed to always keep me honest.

The friendships that one develops are conceivably the most important part of ones graduate education. This is absolutely the case for me. Without them I simply would not have made it. You have all provided me with excellent council and wonderful hugs. You have read my drafts, laughed at me, played with me, visited me in the hospital, drank heroic volumes of wine with me, put my protein on ice in the middle of the weekend, gone for 2AM coffee to push arrows, explained a reaction for the 3rd time, held my hand, listened to me rant and rave about nothing really all that important, and played Ultimate,.. to name only a few things. I want to thank: an excellent friend and room mate Andrew Bogan who put up with for a long time and still doesn't mind having me around on occasion; the beautiful, intelligent, witty, and now a mommy Jamie Moore who had to suffer through a lot of the same experimental hurdles with me; class mates, in particular, Ha Nguyen, Nikki Clegg, Doron Greenbaum, Walley Novak and Kathleen Pendola Novak; the forever smooth Zach Serber; a man who will surely be famous Kurt Thorn; Todd Prey (T-dog); Wesley Straub; the giggling Wah-Wah sisters; everyone in the Sunday 11:59AM ultimate group; Miyuki Susawa (Su-san) for holding my hand through Tokyo University; and Kinkead Reiling who always takes first billing. It is wonderful to look at all of us now and see where we are going. Keep in touch and keep on smiling. It's contagious.

Novel Selective Proteomimetic Inhibitors

By

Timothy R. Geistlinger

Abstract

Department of Chemistry and Chemical Biology, University of California at San Francisco, Genentech Hall, Mission Bay, 600 16th St., San Francisco, CA 94143-2280

Abstract: Nuclear receptor (NR) super-family of signaling proteins is responsible for the integration of physiological signals to maintain the dynamic homeostasis of an organism by regulating growth, development, and differentiation. This signaling network is composed of a series of highly conserved and mutually shared protein-protein interactions that permit the formation and disassembly of macromolecular protein transcriptional regulatory complexes composed of NR, coactivators, corepressors and transcription factors. Specific direct competitive inhibitors of these protein-protein interactions would aid in our understanding of the specific role of each of these interactions in the regulatory complex formation. We chose to target the coactivator interaction with the NR which has been identified to be a critical step in the activation complex formation following ligand binding. Coactivators bind receptors in a ligand dependent manner via the highly conserved alpha-helix LXXLL NR box motif. By mimicking this protein interface we have rationally designed chemical probes that are specific to different NR which inhibit coactivator binding despite the shared binding of the LXXLL NR box.

Here we present the design, synthesis, evaluation and selection of a series of inhibitors that are specific to human Estrogen Receptor Alpha (hERα), human Estrogen Receptor Beta (hERβ) and human Thyroid Hormone Receptor Beta (hTRβ). Additionally we have identified a set of inhibitors that are able to discriminate between three estrogens, 17β-estradiol, genistein (Gen), and an estrogen analogue diethylstilbeserol (DES), with both ERα and ERβ. These inhibitors have identified sub-site specific differences between several NR and their liganded states that have not been identified through traditional

biochemical methods. Our 2.3 Å resolution crystal structures of these inhibitors bound to ERα are providing insights behind the selectivity between receptors at this interface which are generating a new framework for drug design at the protein-protein interactions. We hope to utilize these probes to understand the role of this interaction in NR gene regulation and to investigate the therapeutic potential of targeting this interface in NR related disorders such as breast cancer. We also believe these results will play a seminal role in the identification of novel approaches to therapeutic intervention by characterizing protein interfaces in a variety of signaling networks for target validation and specific inhibitor design.

Table of Contents

List of Figures

List of Tables

Copyright Permission

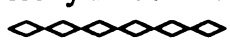
From: "Somers, Holly (CELL)" <hsomers@cell.com>
To: "Tim Geistlinger" <tgeistl@itsa.ucsf.edu>
Cc: "Mountain, Victoria (Cell)" <VMountain@cell.com>
Subject: RE: Thank you and Copyright permission
Date: Wed, 25 Feb 2004 08:48:15 -0600

Hi Tim,

You're very welcome, of course, and I'm sending a few covers to you today.

About permissions: you're free to use the material for your dissertation under the terms of the copyright agreement, but if it eventually is intended for publication elsewhere, you'll need to get in touch with us so we can point you to our permissions office.

Best wishes,
Holly

Holly D. Somers

Managing Editor
Chemistry & Biology
Cell Press
1100 Massachusetts Avenue
Cambridge, MA 02138
(617) 397-2856 phone
(617) 397-2820 fax
hsomers@cell.com

ACS PUBLICATIONS DIVISION GUIDELINES

FOR THESES AND DISSERTATIONS

ATTENTION: STUDENTS, STUDENT ADVISORS, AND TEACHERS

Permission is automatically granted to include your paper(s) or portions of your paper(s) in your thesis; please pay special attention to the implications paragraph below. The Copyright Subcommittee of the Joint Board/Council Committees on Publications approved the following:

Copyright permission for published and submitted material from theses and dissertations
ACS extends blanket permission to students to include in their theses and dissertations their own articles, or portions thereof, that have been published in ACS journals or submitted to ACS journals for publication, provided that the ACS copyright credit line is noted on the appropriate page(s).

Publishing implications of electronic publication of theses and dissertation material

Students and their mentors should be aware that posting of theses and dissertation material on the Web prior to submission of material from that thesis or dissertation to an ACS journal may affect publication in that journal. Whether Web posting is considered prior publication may be evaluated on a case-by-case basis by the journal's editor. If an ACS journal editor considers Web posting to be "prior publication", the paper will not be accepted for publication in that journal. If you intend to submit your unpublished paper to ACS for publication, check with the appropriate editor prior to posting your manuscript electronically.

If your paper has not yet been published by ACS, we have no objection to your including the text or portions of the text in your thesis/dissertation in **print and microfilm formats**; please note, however, that electronic distribution or Web posting of the unpublished paper as part of your thesis in electronic formats might jeopardize publication of your paper by ACS. Please print the following credit line on the first page of your article: "Reproduced (or 'Reproduced in part') with permission from [JOURNAL NAME], in press (or 'submitted for publication'). Unpublished work copyright [CURRENT YEAR] American Chemical Society." Include appropriate information.

If your paper has already been published by ACS and you want to include the text or portions of the text in your thesis/dissertation in **print or microfilm formats**, please print the ACS copyright credit line on the first page of your article: "Reproduced (or 'Reproduced in part') with permission from [FULL REFERENCE CITATION.] Copyright [YEAR] American Chemical Society." Include appropriate information.

Note: If you plan to submit your thesis to UMI or to another dissertation distributor, you should not include the unpublished ACS paper in your thesis if the thesis will be disseminated electronically, until ACS has published your paper. After publication of the paper by ACS, you may release the entire thesis (**not the individual ACS article by itself**) for electronic dissemination; ACS's copyright credit line should be printed on the first page of the ACS paper.

SUMMARY: The inclusion of your ACS unpublished or published manuscript is permitted in your thesis in print and microfilm formats. If ACS has published your paper you may include the manuscript in your thesis on an intranet that is not publicly available. Your ACS article cannot be posted electronically on a publicly available medium, such as but not limited to, electronic archives, Internet, intranet, library server, etc. The only material from your paper that can be posted on a public electronic medium is the article abstract, figures, and tables and you may link to the article's DOI.

Questions? Please contact the ACS Publications Division Copyright Office at copyright@acs.org or at 202-872-4368.

August 1998, March 2003, October 200

List of Publications

Timothy R. Geistlinger.; Andrea McReynolds,; R. Kiplin Guy, Specific Inhibitors Select for Different Ligands with Each Estrogen Receptor Isoform, *Chemistry and Biology*; **2004**; (11) pp 273-281.

Timothy R. Geistlinger. R. Kiplin Guy, Novel Selective Inhibitors of the Interaction of Individual Nuclear Hormone Receptors and a Mutually Shared - Steroid Receptor Coactivator 2, *J. Am. Chem. Soc.*; **2003;** *125*(23) pp 6852 – 6853

Timothy R. Geistlinger. R. Kiplin Guy, Steroid Receptor Coactivator Proteomimetics, *Methods In Enzymology*, **2003**, Vol. 364, Chapter 13, pp 223-246

Timothy R. Geistlinger. R. Kiplin Guy, An Inhibitor of the Interaction of Thyroid Hormone Receptor β and the Glucocorticoid Receptor Interacting Protein1, *J Am Chem Soc* **2001**, *123*, 1525-1526

Chapter I - An Introduction – The Nuclear Receptor Biology And The Coregulator Interface

Summary

The nuclear receptor superfamily is responsible for integrating biological signals into the precise regulation of many important developmental and physiological processes. These receptors are highly conserved intracellular transcription factors that change the expression of hormone-responsive target genes.[1-4] Missintegration of signals by the nuclear receptor signaling networks has been correlated with a broad range of common pathologies including metabolic diseases, breast and prostate cancer, arthritis, obesity, and diabetes, advancing the nuclear receptors (NR) to become one of the top targets of basic and pharmaceutical research.[5] Current drug design efforts have been focused on the development of synthetic ligand antagonists which compete with natural hormones to modulate the nuclear receptors transcription regulation. These drugs have verified the nuclear receptors as a therapeutic target in the clinic with treatments for breast cancer, however, it has become increasingly difficult to determine how these agents modulate activity. Additionally, with the exception of a few examples, it has been difficult to separate their tissue dependent beneficial effects from the pathological activities associated with these receptors. Therefore, while many antagonists are valuable therapeutic agents in the treatment of hormone-dependent diseases, their use is often associated with unwanted side effects. Identifying an alternative approach to regulating the nuclear receptor signaling network has consequently become of increasing interest.

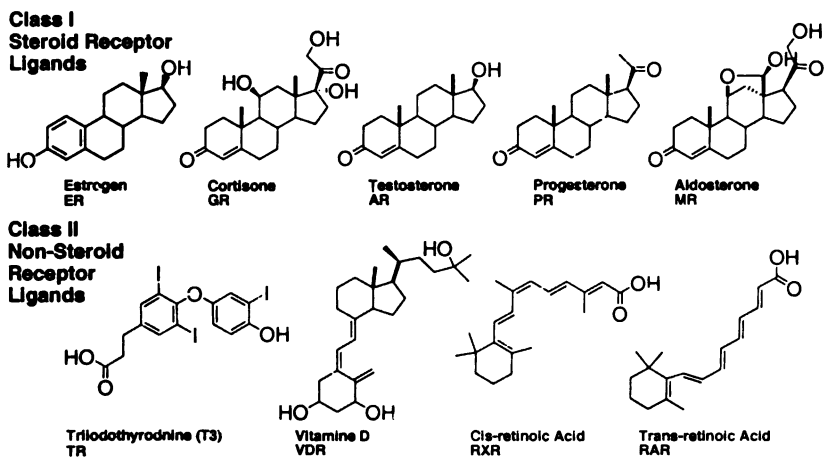


Figure 1-1. Nuclear Receptor Ligands

Nuclear receptors have traditionally been placed into three categories based on their ligand binding: 1) Class I - Steroid Receptors and their ligands – Estrogen Receptor (ER), Glucocorticoid Receptor (GR), Androgen Receptor (AR), Progesterone receptor (PR), Mineralcorticoid Receptor (MR); 2) Class II – Hormone receptors – Thyroid Hormone Receptor (TR) binds Triiodothyronine (T_3), Vitamine D Receptor (VDR), Cis-Retinoic Acid receptor (RXR), and the all Trans-retinoic Acid Receptor (RAR); and 3) the Orphan Receptors for which the cognate ligands have yet to be clearly identified.

The molecular mechanisms by which nuclear receptors regulate transcription in a ligand dependent manner was recently elucidated to reveal that hormone binding regulates the interaction of nuclear receptors with coregulators, coactivators or corepressors, which respectively amplify or reduce the expression of the hormone-responsive genes above or below that of the basal level of transcription. Upon hormone binding a small hydrophic cleft on the nuclear receptor was exposed to recognize and bind a small, α-helical, three leucine repeat (LXXLL) –where X indicates any amino acid - motif on coactivators[1,4] to form a protein-protein interaction which initiates the formation of a macromolecular protein transcription activation complex.[6] Biochemical studies subsequently revealed that this site could be blocked by peptides containing the LXXLL motif and recombinant peptides could inhibit the interaction *in vitro* and inhibit transactivation in cells, respectively.[7,8] These results suggested a possible hormone-independent mechanism for regulating the transcriptional activity of nuclear receptors. These discoveries moved the spotlight onto targeting the nuclear receptor•steroid receptor coactivator (NR•SRC) protein-protein interaction for novel drug design. In order to be useful, however, the inhibitors must be able to not only effectively inhibit the interaction but also to be selective for one nuclear receptor within the superfamily of the more than 50, highly conserved, NR which all seem to interact with the LXXLL motifs. Additionally, in order to be potentially useful as a therapeutic, the basis for selectivity must be reduced to a size that would be achieved in a small molecule.

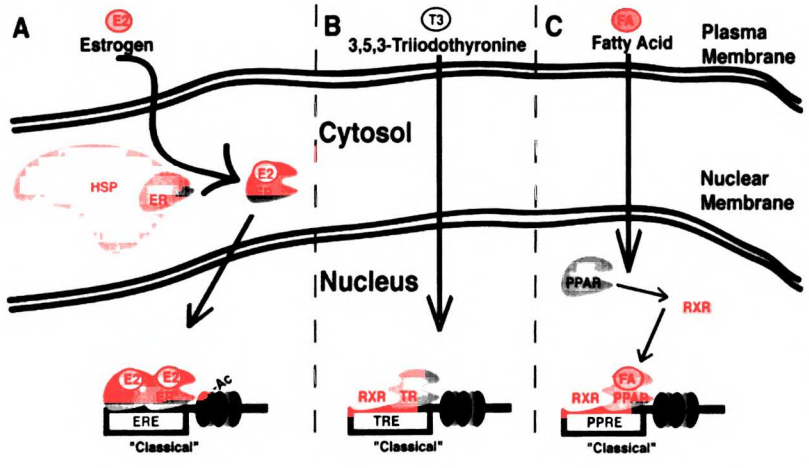


Figure 1-2 The NR - Intracellular Receptors

Nuclear receptors have different mechanisms of action. **A**) Estrogen Receptors (ER) as well as many of the steroid receptors, exist in the cytosolic space bound to heat shock proteins (HSP90). The receptor undergoes a conformational change upon ligand binding to release the HSP and localize to the nucleus. ER then binds to a target DNA estrogen response element (ERE) as a homo-dimmer. Dimerization and DNA binding are ligand dependent. **B**) TR and many of the Type II receptors exist primarily in the nucleus, bound to a specific response element primarily as a

heterodimmer with RXR but may bind as a homo-dimmer depending on the response element. The ligand T_3 binds TR to create a conformational change at the thyroid hormone response element (TRE) but does not determine the dimerization state or DNA binding state of the receptor. **C)** Other receptors, like the peroxisome proliferating activated receptor (PPAR) seems to exist free in the nucleus. Upon binding ligand, a fatty acid molecule (FA), the receptor dimerizes with RXR and binds the DNA peroxisome proliferating activated receptor response element (PPRE). Unlike TR, this receptor seems to bind DNA in a ligand dependent manner.

Nuclear Receptor Signaling

Nuclear receptors function to integrate multiple signals in the form of growth factor signaling cascades, hormones, and steroids (Figure 1-1) to precisely coordinate genomic expression. The mechanism of action for each NR is different (Figure 1-2), however, using the estrogen receptor as an example, in the absence of ligand ER is cytosolic and bound by a chaperone protein such as the 90 kD heat shock protein (Hsp90) chaperone. Upon binding estrogen the receptor releases the Hsp90 and localizes to the nucleus and binds particular DNA response elements, as a homo-dimer. At this point the receptor may stay on the DNA with or without ligand in the dimerized state. Functioning as a molecular hub, NR coordinate the binding of a collection of coregulatory proteins and transcription factors, continually assembling and disassembling, different macromolecular complexes at these specific response elements, on the seconds timescale (Figure 1-3). Here, the NR function as a molecular switch within a signaling network to determine the composition of the complex and promote three levels of gene expression: (i) a ligand dependent trans-activation or AF-2 pathway; (ii) a ligand independent basal AF-1 activation, and (iii) a transrepression level where gene expression is actively repressed. Pressure from any direction, such as one of the afore mentioned signals, a change in the expression levels of particular coregulatory proteins, or the posttranslationally modified state of any of the proteins, seems to alter the composition of these complexes and modulate the regulation of the target gene. The overexpression of a particular coregulator seems to be capable of driving a pathway, simply through mass action. For example, several pathologies have been directly correlated to an overexpression of the coactivators themselves.[9,10] Depending on the physiological context and the relative concentrations of coregulators for example, the ligand effect can change from functioning as an agonist in one tissue to an antagonist in another. Each step in this network is in equilibrium and can be modulated by any one of a number of events to shift

the reaction in one direction or another, Le Chatlier's principle applied to macromolecular biology.

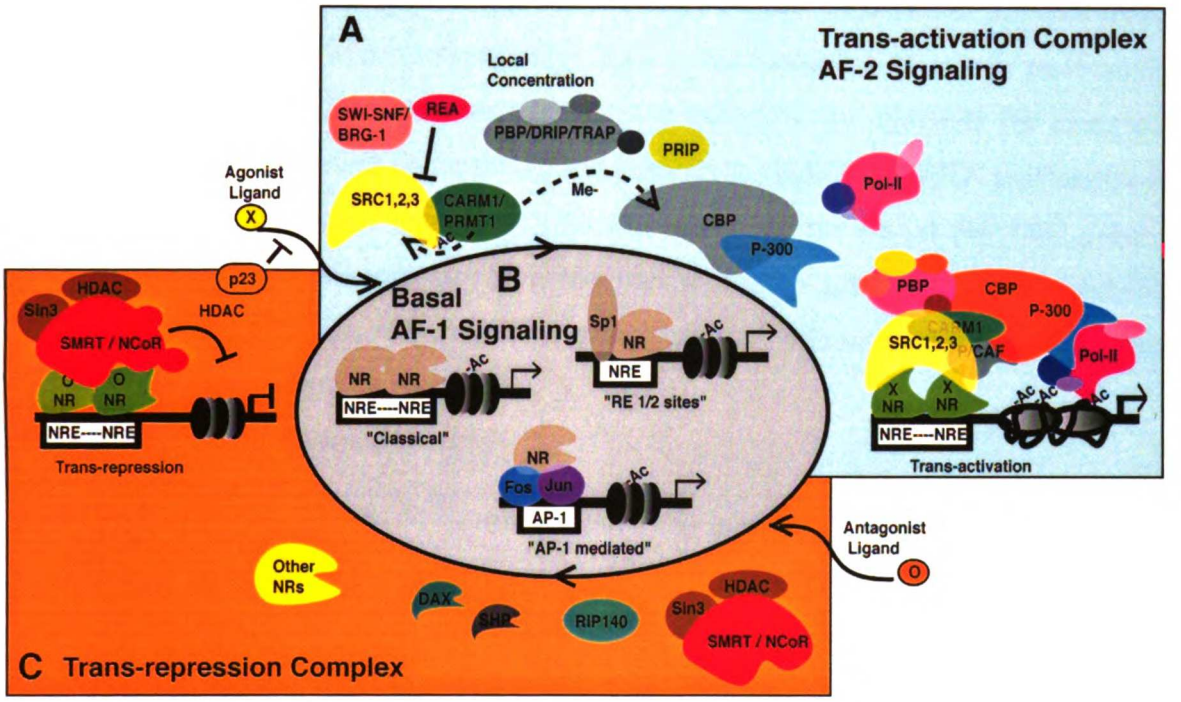


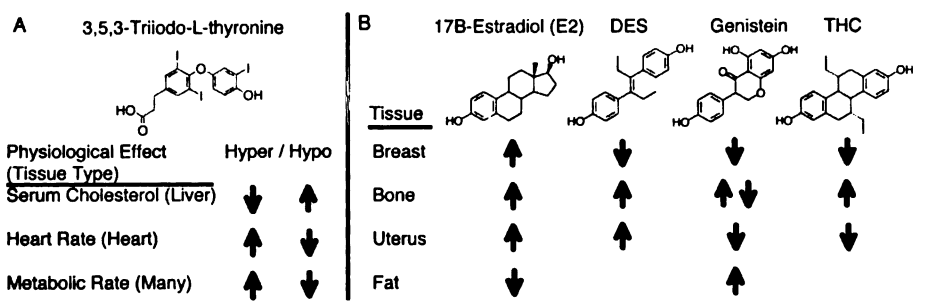
Figure 1-3. NR Network of Combinatorial Signaling Equilibrium

The nuclear receptors (NR) function as soluble, intracellular, transcription factors which act as a "hub" to bind specific nuclear response elements (NRE) as "classical" homo- or hetero-dimers, as monomers to response element half sites "RE 1/2 sites" or as a tethered complex via other proteins like that of the Fos and Jun AP-1 complex "AP-1 mediated". Depending on the physiological context, which type of response element, or the ligand binding state, will determine what the role of the NR is at that target gene. The NR also function as a molecular switch. (**A**) In an agonist (X) bound state, at a classical response element, NR are known to coordinate the binding of steroid receptor coactivator proteins, (SRC1/SRC2/SRC3) as the first critical protein-protein interaction. The SRC proteins function to coordinate the binding of a coactivator arginine methyltransferase (CARM) as well as other coactivators (DRIP/TRAP/PBP). Together some of these proteins have the ability to covalently modify other proteins/chromatin as well as recruit the transcription initiation complex CBP/P300, transcription factors and RNA polymerase II (Pol-II). This "Trans-activation Complex" functions to modify the Chromatin/DNA to unwind the DNA and promote transcription of the target gene. (**B**) A ligand independent, or basal level of gene regulation also exists when the NR are bound at specific response elements. This is an ATP dependent regulation that is known as AF-1, incorporating the SWI-SNF/BRG-1 complex of proteins that are coordinated to the N-terminus of the NR. (**C**) In the absence of ligand, corepressor (SMRT/NcoR – Sin3A) Histone Deacetylase (HDAC) complex of molecules may be coordinated to the NR at specific response elements. The HDAC "Trans-repression Complex" functions by deacetylating the histones which condenses the chromatin structure and inhibits the binding of the polymerase. This causes a down regulation of the target protein. The system is in equilibrium, constantly adjusting to assemble and disassemble the different protein complexes at DNA response elements to properly regulate the gene transcription. Disassembly appears to utilize the 23kD chaperone (p23) which functions to dissociate the NR from the DNA, compete off SRC proteins, and increase the off rate of ligand.[11,12]

In the agonist (estrogen) bound state the receptor goes through a conformational change to permit the association of a family of proteins known as steroid receptor coactivators. This is known to be the primary, ligand dependent, protein-protein interaction (NR:SRC) which subsequently leads to the formation of a large transcription activation complex. This complex functions to acetylate the histones for chromatin decondensation, unwinding of the DNA, and promote the binding of RNA polymerase for transcription of the target gene. Theoretically inhibiting the initial NR:SRC protein-protein interaction in this sequence of events would create a novel approach to inhibiting only the ligand dependent AF-2 transactivation pathway, without antagonizing the basal, AF-1, or the transrepression pathways.

Figure 1-4. Tissue Specific Effects

Ligands seem to act as either agonist or antagonists in a tissue dependent context. (**A**) Hyperthyroidism – high levels of Thyroid Hormone (T_3) functions to decrease levels of serum cholesterol in the liver while increasing the heart rate and metabolic activity in other tissue. Low levels of T_3, Hypothyroidism, has the opposite effects in each of those tissues. cDNA Arrays [13] have shown that thyroid hormone seems to function equally as both an agonist and antagonist increasing and decreasing gene transcription, respectively, within the same tissue. (**B**) Different estrogens such as Estradiol (E_2), Diethylstilbesterol (DES), Genistein (Gen) and the synthetic analogue (THC) each function differently in a tissue (Breast, Bone, Uterus, Fat) dependent manner. Arrows represent an increase or decrease in the rate of tissue growth.



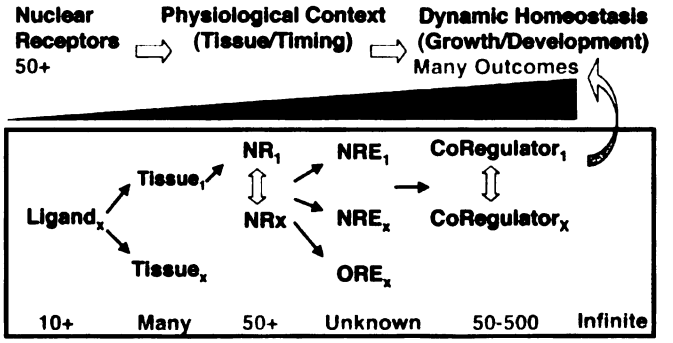
The nuclear receptor has a distinct biological response to different natural and synthetic estrogen-like ligands. [14] [15-17] giving rise to the development of selective estrogen response modulators (SERMs) for breast cancer chemotherapeutics. The effect of the ligand depends primarily on their ability to differentially alter the structure, stability, and bio-availability of the receptor LBD.[18] Why these unique effects take place is not clear. Estrogen, and estrogen-like ligands function by binding with high affinity to the ligand binding domain (LBD) to define the hydrophobic core of the domain. As a result the ligands allosterically define the overall structure creating a conformational change in the LBD to promote or inhibit the binding of coregulator proteins consisting of coactivators and corepressors. Depending on the physiological context, the receptor isoform, and

which coregulators are present, a particular ligand – thyroid hormone for example – may result in an increase in the relative tissue activity or a decrease (Figure 1-4). Likewise, different analogues of estrogen may cause a tissue dependent estrogenic, partial estrogenic, or an anti-estrogenic, response.[19] Until now the focus of ligand allosteric effects has been on gross structural changes like the relative modulation of the helix 12 position within the LBD and its ability to create or conceal the SRC binding pocket to act as an agonist or an antagonist, respectively. However, it is not quite that simple as each ligand seems to have a different effect on the relative equilibrium between these states and between estrogen receptor isoforms.[18,20-22] Such a change can create a wide variety of effects down field which is context dependent.

It is likely that the different ligand effects in different tissues are due to a dissimilarity in the combinatorial arrays of coregulators in each tissue type. A recent theory has proposed that this combinatorial array of coregulator molecules that has created the phenotypic variation in cellular and organism development from such a limited gene pool. This can be extended to explain phenotypic variation in evolutionary biology given a gene pool with a relatively high degree of homology. For example, this combinatorial variation may explain why, organisms - humans vs. humans or humans vs. chimpanzees - have such a high degree of phenotypic variation. In a review by O'malley et.al. [2,3,23] it is suggested that the rates, and timing of gene expression may be tantamount to the genes themselves in manifesting a particular phenotype. Based on our current understanding of the combinatorial pattern of protein-protein interactions and the effect on genomic outcome and cellular biology, a single event within the signaling cascade has exponential effects (Figure 1-5).

Figure 1-5. Nuclear Receptor Combinatorial Biology

The ability of the nuclear receptors to integrate a wide variety of signals to effectively regulate the genomic expression of and organism throughout the growth and development is due to the combinatorial assembly of multiple, context dependent regulatory complexes. There are approximately 50 nuclear receptors NR_x which appear to be ubiquitously expressed ($Tissue_x$), that respond to different ligands (10+), bind to multiple different specific response elements (NRE_x), and share multiple (50-500) Coregulator proteins ($Coregulator_x$). Each of these different combinations leads to a different set of genes being regulated to different

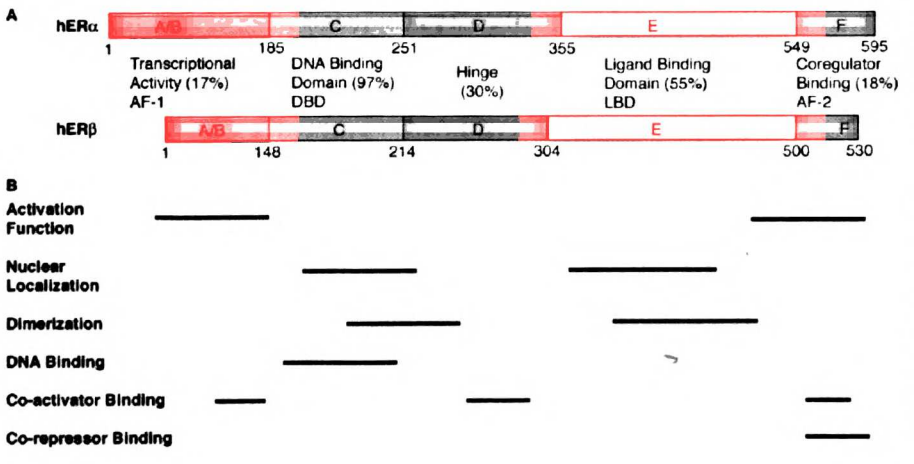


levels in different tissue at different time points.

The Nuclear Receptors

The nuclear receptors are soluble proteins composed of five functional domains (Figure 1-6). The N-terminal A/B domain is known to have ligand independent transcription activation domain forming protein interactions with Fos and Jun as well as the SWI/SNF/BRG-1 proteins and is referred to as the AF-1 domain. The DNA binding domain C functions to bind directly to the DNA response element and has the highest degree of homology between the different NR. A hinge region D connects the DNA binding domain to the Ligand Binding Domain (LBD). The LBD not only binds the cognate ligand, but also serves a role in nuclear localization, dimerization, and coregulator binding AF-2 transcription activity. When considering the multiple functions located within the LBD it is clear how different ligands can affect each of these functionalities differently. For example, different ligands are known to alter the dimerization constant, nuclear localization, and coregulator binding. Alterations in dimerization can have an indirect effect on DNA binding, (important at some response elements and not others). Changes in coregulator binding can select for different sets of coactivators, select for corepressors or alter the NR incapable of coregulator binding to manifest a dominant negative effect. As a result, proper dissection of the functionality or each domain and the individual interfaces would require direct competitive inhibitors of individual functionalities, as opposed to non-competitive allosteric inhibitors that modulate multiple interfaces simultaneously.

Figure 1-6 Map of Nuclear Receptor Primary Structure and Functionality



Like the other members of the nuclear receptor superfamily, the estrogen receptor isoforms are highly homologous proteins of approximately 600 amino acids composed of five domains (**A**). The primary structures of human ERα and ERβ are shown with domains labeled A through F, relative amino

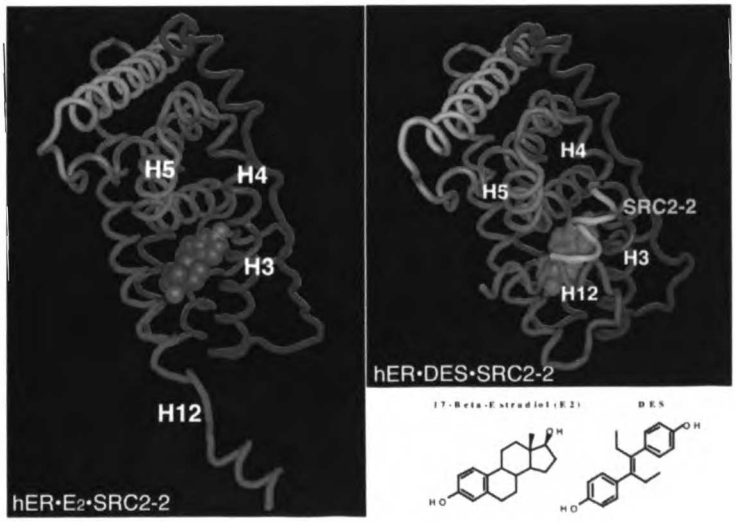
acid position numbers, functional domain labeling and relative percent homology between the two isoforms. (**B**) Functional mapping of NR shows several known points of activation function, nuclear localization, NR dimerization, DNA binding, and coregulator protein-protein interactions, of the nuclear receptors which are not distinct, but rather overlap in a way that is difficult to isolate.

Ligand Dependent Conformational Change

The ligand binding domains (LBDs) of NRs adopt distinct conformations as aporeceptors, agonist-bound, or antagonist-bound species to support or preclude interactions with proteins such as chaperones, corepressors, or coactivators [24]. LBDs remain functional as compact `molecular switches.' Structural analyses of the LBDs from receptors for retinoids (RXR) [25,26] and RAR[27], thyroid hormone (TR),[28] estrogen (ER),[29] and progesterone (PR),[30] revealed a common overall fold for the domain, despite substantial sequence divergence. The structure of the NR LBD consists of 12 α-helices and 4 β-strands organized in three layers.[28] Structural studies revealed that the nuclear receptor LBD underwent a structural change upon ligand binding in the twelfth α-helix (Figure 1-7) within the LBD which acts as an allosteric switch on the surface of the protein.[18] In the case of an agonist ligand, this helix moves to reveal a small hydrophobic pocket between helicies three, four, five and twelve which binds three leucines within an amphipathic α-helix LXXLL motif of the coactivator. Conversely an antagonist ligand modulates this helix-12 to conceal the pocket or alter it in a way to bind an extended IXXLXXII α-helix of a corepressor. In the absence of hormone the helix-12 presumably fluctuates between these states as an ensemble of conformations. Different ligands are known to modulate this equilibrium between these states.[21]

Figure 1-7. Estrogen Receptor Ligand Binding Domain Crystal Structures

here different crystallographic structure solutions are shown of the human estrogen receptor alpha (hERα) ligand binding domain (LBD). The LBD of ERα, like other NR, is composed of twelve α-helicies (H1-H12) with a pocket in the hydrophobic core of the receptor. (Left) shows the crystal structure of hERα•E2 (PDB:1A52). Subsequent to ligand binding the helix-12 (red) of the domain wraps up and over the ligand to bury it in the core of the domain as shown



in the right panel. (Right) hERα•DES•SRC2-2 (PDB:3ERB) Helix-12 comes up to create a pocket on the surface of the receptor where the α-helix of the SRC2 LXXLL NR box 2 is shown binding (yellow). This pocket is composed of helicies H3-H5 as labeled.

The Coactivators

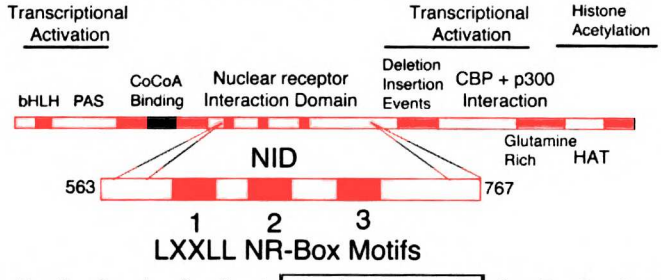
Ligand dependent NR signaling requires direct interaction between NR and the steroid receptor coactivators (SRC). Agonist-bound LBDs can bind coactivators from the p160 family, which includes at least three distinct members, SRC-1 [31]; also NcoA-1 [32], p/CIP [33]; also AIB1 [10], TRAM-1 [34], RAC3, ACTR [35], and TIF2 [36]; also GRIP1 [37,38], NcoA-2.[24,39] The SRC's have been reduced to three sets based on their genes: SRC1, SRC2/GRIP1/TIF2, and SRC3/ACTR/Rac3/pCIP/AIB1.[10,36-38,40-44] A unified nomenclature, in which the proteins are referred to as SRC1, SRC2, and SRC3, has been proposed by O'Malley and is adopted in this dissertation. These ~160-kD coregulators seem to have multiple domains with several different functions (Figure 1-8). The SRC's themselves have several well-characterized activation domains: the amino-terminal bHLH and PAS (Per/Arnt/Sim) domains; Coiled-Coil Coactivator (CoCoA) interaction motif [45,46], one that recruits CBP/p300;[47] another that recruits an arginine methyltransferase CARM;[48] and a domain with histone acetyl transferase activity. Homozygous disruption of the genes encoding each SRC causes fairly distinct phenotypes in mice thus suggesting the possibility of distinct mechanisms or activities for each SRC. [49,50] However, there is evidence of compensatory upregulation of SRC2 following disruption of SRC1, indicating the possibility of partially overlapping function. The underlying biochemical mechanisms that may allow for conservation or separation of SRC function in ligand dependent signaling are poorly understood. Inhibitors that would allow selective disruption of particular NR•SRC interactions *in vivo* would be very useful in elucidating such mechanisms.

Subsequent detailed biochemical investigations of SRC proteins and their function in NR mediated ligand dependent transcription revealed the nuclear receptor interacting domain (NID) of coactivators, which includes multiple interaction motifs known as NR boxes with a consensus sequence of $L_1XXL_2L_3$. Each NR box of the SRC's NID can have different affinities for a particular NR•ligand•promotor triad [51] and can bind in a cooperative[52] or non-cooperative[7,53,54] manner. Despite recent findings of additional interactions of SRC with NR outside the NID,[55] blocking the interaction of NR and SRC

in cellular models by overexpression of fusions of SRC NR box peptides inhibits gene transcription from elements normally responsive to hormone induced signaling.[8] In fact, short NR box peptides and α-helical proteomimetics containing the NR box sequence can effectively compete with the entire NID of the SRC2.[7] This suggests that blocking the interaction at this site alone would be suitable for inhibiting the entire protein-protein interaction.

Figure 1-8. The SRC p160 proteins & NR box sequences

These ~160-kD proteins include amino-terminal bHLH, and PAS domain, a CoCoA interacting region[56-58], domains for interaction with nuclear receptors (NR interaction domain, NID), CREB-binding protein (CBP), and p300, several regions with transcriptional activation potential, c-terminal histone acetylation (HAT) potential. [36,58-61]. In p160 coactivators, and in unrelated coactivators such as TIF1 or RIP140, interaction with NRs is mediated by multiple LxxLL motifs, each residing within distinct patches of conserved sequence termed NR boxes [52,58,62-66].



NR-Box	-7	-6	-5	-4	-3	-2	-1	1	2	3	4	5	6	7	8	9	10	11
SRC1-1	Y	S	Q	T	S	H	K	L	L	Q	L	L	T	T	T	A	E	Q
SRC1-2	L	T	E	R	H	K	I	L	H	R	L	L	Q	E	G	S	P	S
SRC1-3	E	S	K	D	H	Q	L	L	R	Y	L	L	D	K	D	E	K	D
SRC2-1	D	S	K	G	Q	T	K	I	L	Q	I	L	T	T	K	S	E	Q
SRC2-2	L	K	E	K	H	K	I	L	H	R	L	L	Q	D	S	S	S	P
SRC2-3	K	K	K	E	N	A	L	L	R	Y	L	L	D	K	D	D	T	K
SRC3-1	E	S	K	G	H	K	K	L	L	Q	L	L	T	C	S	S	E	D
SRC3-2	L	Q	E	K	H	R	I	L	H	K	L	L	Q	N	G	N	S	P
SRC3-3	K	K	E	N	N	A	L	L	R	Y	L	L	D	K	D	D	P	S
								L	X	X	L	L						

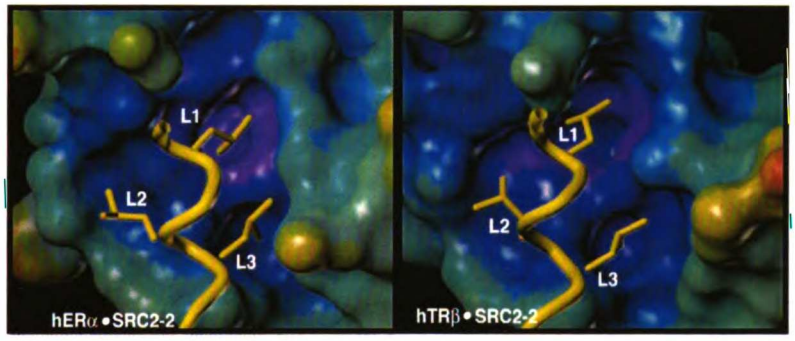
Many factors around NR•SRC regulation and function remain poorly elucidated. How one nuclear receptor selects for a particular SRC sequence is not clear. It is also well known that many NR seem to share multiple SRCs and do so by binding the same set of NR boxes. For example, hERα, hERβ, and hTRβ share the use of SRC2, despite the fact that they regulate entirely different gene transcription pathways.[65,67] *In vitro*, hTRβ and both ER isoforms tightly bind the second NR box of SRC2 (SRC2-2), ^{685}EKHKI**L$_1$**ER**L$_2$L$_3$**KDS^{697}, in the presence of their native ligands thyroid hormone (T_3) and estradiol (E_2), respectively. The ER isoforms also interact with SRC2-1 with lesser affinity while the TR's interact with SRC2-3. Biochemical studies employing phage display and recombinant peptide libraries have determined that the three leucines within the LXXLL motif are a requisite for binding and that the sequences surrounding the motif may very. Those flanking sequences seem to impart some selectivity for binding to different NR. An analogous motif (I/LxxII) has been identified for corepressors SMRT and NCOR and the binding sites for coactivator and corepressor may be partially

overlapping.[65,67-70] And finally, recent functional studies of SRC proteins have begun to break down the coactivator paradigm by demonstrating that SRC 2 can act as a corepressor in an NR and NR response element dependent fashion.[71] Corepressor activity of SRC 2 was mapped to a domain distinct from the two known SRC 2 activation domains. This function was not detected in the other two p160 family members, SRC 1 and SRC 3. The underlying biochemical mechanisms that may allow for conservation or separation of function in ligand dependent signaling are poorly understood. Inhibitors that would allow selective disruption of particular SRC•NR interactions *in vivo* would be very useful in elucidating such mechanisms.

The Nuclear Receptor : Coactivator Interface

Figure 1-9. The Nuclear Receptor Interface NR•SRC

The SRC2-2 peptide (EKHKILHRLLQDS) LXXLL motif is shown binding as a α-helix to a shallow, hydrophobic groove on both the Estrogen receptor (hER•DES•SRC2-2 – PDB:3ERB) and the thyroid hormone receptor (hTRβ•T_3•SRC2-2 – PDB:1BSX) in the presence of agonist ligand. Efficient interaction within both the ER•SRC and TR•SRC interfaces rely on both the hydrophobicity of the LxxLL motif and the stereochemical properties of each of the three leucine side chains L1, L2, and L3. Receptor conolly surfaces were generated with an electrostatic gradient (*Sybyl*) to demonstrate the degree of similarity.



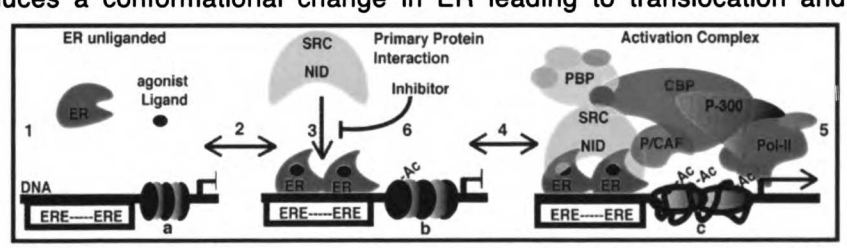
Structurally, the interactions of hERα and hTRβ with the SRC2-2 peptide are very similar to each other (Figure 1-9) as well as other NR•Coactivator structures solved to date.[7,18,65,67,72-75] Each seems to interact through similar surfaces with a shallow hydrophobic groove on the NR surface binding to an induced fit, amphipathic, α-helical motif on the SRC2-2 NR box, burying the conserved three leucines on the hydrophobic face of the NR box helix. While each of the leucines in the NR box is critical to the interaction,[61] *in vivo* specificity in recruiting a particular SRC appears to be induced by the sequences immediately flanking the NR boxes[8,64,75,76] rather than the geometry of the leucine side chains. Manipulation of NR box peptide sequence outside of the conserved $L_1XXL_2L_3$ motif has afforded selective peptide inhibitors of the interaction of particular NR and SRC's, presumably by taking advantage of these extended interactions.[8,77,78]

However, such extended elements would not be available to a small molecule disrupting the interaction and it is therefore uncertain how such findings reflect upon the development of such a drug.

Each hydrophobic leucine within the LXXLL NR box α-helix plays a critical role in the formation of the NR•SRC complex. *In vitro* and *in vivo* experiments showing that substitution of multiple leucine residues of the motif by alanine abrogated physical and functional interactions between p160 family coactivators and NR. Additionally, replacement of single leucine residues of SRC2-2 by phenylalanine reduced competition by the mutant peptides for TRβ LBD by 60- to 100-fold relative to wild-type peptide while replacement of the isoleucine resulted in a ~10-fold reduction. This suggested that efficient interaction within this interface relies on both the hydrophobicity of the $L_1XXL_2L_3$ motif and on stereochemical properties of the leucine side chains. However, nature is limited to only four such hydrophobic residues, Leucine, Isoleucine, Valine, and Phenylalanine with a narrow window of structural and conformational diversity. Also, these different residues may play an equally important role in the secondary structure and propensity of helix formation, a factor that would dramatically modulate the interaction. Therefore, we hypothesized that by utilizing non-natural residues and a rigid α-helical scaffold it would be possible to select for sub-site differences within the interfaces between different NR.

Figure 1-10. Targeted inhibition of Ligand Dependent Nuclear Receptor Regulatory Transcription Complex Assembly.

Schematic model of the dynamic assembly of the transcription activation complex by agonist bound estrogen receptor (ER) and function of steroid receptor coactivator (SRC) binding inhibitors. **1**) In the absence of ligand, chromatin is unmodified (a) and transcription at the DNA estrogen response element (ERE) is unaltered. **2**) Binding of agonist ligand to the ligand binding domain (LBD) of ER induces a conformational change in ER leading to translocation and homodimerization on the ERE. **3**) Liganded ER on the ERE recruits SRC's using the NR box ($L_1XXL_2L_3$) of the nuclear receptor interacting domain (NID). **4**) Subsequently, the ER•SRC complex recruits other proteins to form the activation complex where chromatin is modified (b) and **5**) transcription of the ERE gene commences. **6**) Direct competitive inhibition of SRC binding to NR will block the initial step (3) of activation complex formation and thus prevent transcription. Such an approach should be ligand dependent and clarify the role of this interface in the formation of the transactivation complex,



Inhibiting the NR•SRC Interaction

The NR•SRC interface is an excellent point of intervention to dissect this complex. and understand the specific role of this interface in selected NR signaling networks. (Figure 1-10). However, several fundamental properties of protein-protein interactions – like that of the NR•SRC interaction - have hindered development of drugs directed at these targets. These hindrances include a paucity of lead compounds, large interface sizes that seem to demand "large molecules", high degrees of structural plasticity in the targets that hinder rational drug design, and topological simplicity of the often highly conserved interfaces that leave medicinal chemists few handles for building selective drugs. Due to the relatively small volume and rigid character generally desired in drugs, the latter two characteristics conspire to make the production of receptor isoform selective compounds particularly difficult. Similarly, the NR•SRC protein-protein interaction is highly conserved across the NR superfamily, all the SRC family of proteins utilize the same $L_1XXL_2L_3$ binding motif with three leucines, and these coactivators seem to be shared by most or all the NR proteins. In order to effectively target one NR and its interaction with the coactivator proteins, it is vital that these inhibitors are not only selective but that this selectivity is identified within this small hydrophobic pocket which could be effectively translated into a small molecule.

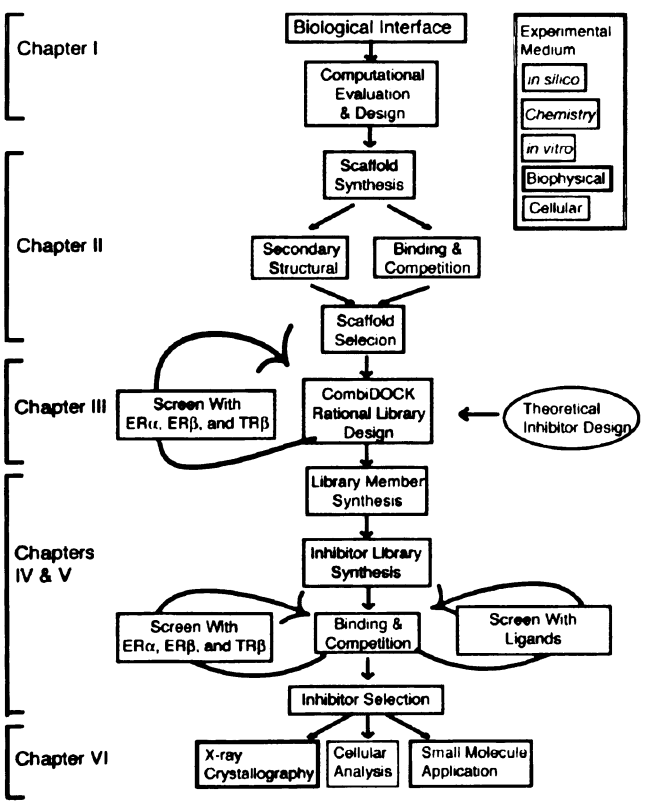


Figure 1-11 Flow Chart of Scientific Approach

Based on computational design *in silico* (blue frame) we designed and synthesize (Green frame) a series of compounds for *in vitro* testing (framed in red). A select set of inhibitors were then subjected to biophysical analysis (Black frame) by x-ray crystallography, cellular analysis (light blue) and small molecule application.

Inhibitor Design

We therefore rationally designed a series of SRC2-2 NR box LXXLL mimetic inhibitors [79-83] with one of three leucines substituted on an α-helical constrained SRC2-2 proteomimetic scaffold.[81] The identification of selectivity within this relatively small motif, unlike other methods that flank this region could be effectively mimicked with a small molecule inhibitor. The crux, however, is identifying significant differences within this relatively small, highly conserved, hydrophobic pocket which act as handles for a small molecule.

To explore the possibility of targeting this interface we decided to utilize a rational combinatorial chemistry design and combine computational power with the harnessed serendipity of combinatorial chemistry. Based on computational evaluation we hypothesized that there were significant differences in the shape and charge distribution in the conserved, hydrophobic groove of nuclear receptors which could be sufficient enough to allow the development of specific inhibitors. Through scaffold designs, *in silico* screening, and rationally designed chemistry, we were able to synthesize a library of 87 inhibitors with single substitutions at each leucine position and through *in vitro* binding studies select for inhibitors that were specific to different nuclear receptors, different receptor isoforms, as well as select between different ligand bound states of the receptors.

The scientific approach we created to address this hypothesis as outlined in the subsequent chapters (Figure 1-11) required several firsts: the design and synthesis of an effective scaffold (Chapter II); the computational analysis of different NR hydrophobic SRC recognition pockets for rational design of a combinatorial library (Chapter III); and the identification of selective inhibitors (Chapters IV and V). The results showed that while we were able to identify selective inhibitors the basis for the selectivity is not entirely clear. Sub-site differences within the pockets between each receptor may or may not be contributing to this selectivity but there is no strong structure activity relationship for each of the receptors or the inhibitors. Our hypothesis was that the binding mode of these inhibitors on the surface of the receptors could be different from that of the native crystal structures. We proposed that either the inhibitors were binding in a slightly different orientation and/or that the receptor surfaces were plastic and changing in

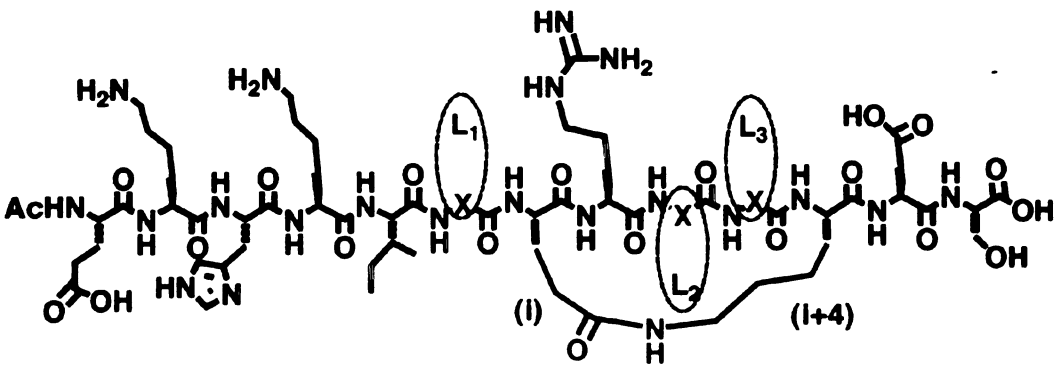
response to the different inhibitors. Only through crystallographic evaluation (Chapter VI) of the interaction between ERα and the inhibitors can we elucidate the mode of binding and possibly the basis for the selectivity we have observed.

Finally, we believe this approach will be useful in generating novel inhibitors of protein-protein interactions for exploring the role of specific interfaces within biological signaling networks, for target validation, for potential therapeutic intervention, and to determine the feasibility of small molecule drug design. In the case of the nuclear receptor signaling cascade this methodology revealed not only subtleties of the energetics and sub-site differences within this interface, but also its role in selecting for different coactivators, and the potential to bind sequences that are different from that of the LXXLL motif. As this knowledge evolves, the ability to inhibit these interactions specifically will pave the road to novel development of selective nuclear receptor-regulating chemotherapeutics and more refined therapies for hormone-dependent diseases.

Chapter II - Selecting a Stable SRC Proteomimetic Scaffold

Timothy R. Geistlinger; R. Kiplin Guy., An Inhibitor of the Interaction of Thyroid Hormone Receptor β and the Glucocorticoid Receptor Interacting Protein1, *J Am Chem Soc* **2001**, *123*, 1525-1526

Timothy R. Geistlinger.; R. Kiplin Guy, Steroid Receptor Coactivator Proteomimetics, *Methods In Enzymology*, **2003**, Vol. 364, Chapter 13, pp 223-246

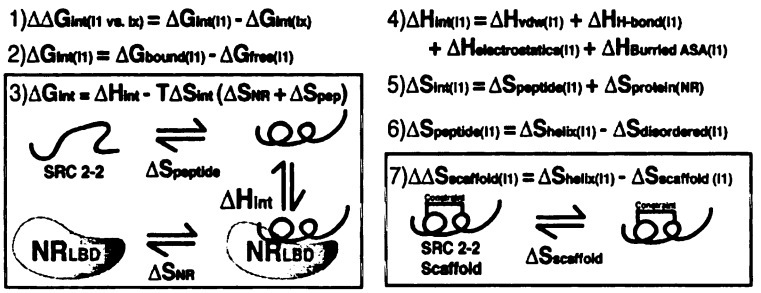


Introduction

In order to proceed with our plan to rationally design a combinatorial library of SRC2-2 LXXLL motif proteomimetics it was necessary to first identify an effective scaffold for the presentation of the leucine mimetic library members. Three criteria were established. The scaffold must: 1) constrain the LXXLL motif into the conformation of an α-helix in solution; 2) present the LXXLL motif to non-selectively bind and compete for the coactivator binding pocket on the nuclear receptor; 3) be stable; and 4) be amenable to solid phase, high throughput synthesis. Because we ultimately wanted to measure the change in the free energy of the interaction $\Delta\Delta G_{int}$ between compounds as a function of the enthalpic contribution $\Delta\Delta H_{int}$ from each leucine substitution in interaction it was necessary to minimize the variability due to the entropic cost of helix formation $\Delta\Delta S_{peptide}$(Figure 2-1).

Figure 2-1. Induced Helical Interface - Free Energy, Enthalpy vs. Entropy

Relative changes in the free energy $\Delta\Delta G_{int}$ of a protein-protein or protein-inhibitor interaction (eq.1,2) can be represented as a function (eq.3) of enthalpic ΔH_{int} and entropic ΔS_{int} changes in and between the two interacting species - Inhibitor(I) and NR. Each of these terms can be broken down into separate components based on their solution vs. binding states (free vs. bound). It is the relative difference between these states that determines the ability of an inhibitor to effectively bind and compete for a target protein. The enthalpic term for any of these states can be reduced to changes involving van der Waals interactions, hydrogen bonding, electrostatics and buried hydrophobic accessible surface area (eq. 4). Entropy may be reduced to changes in the degrees of freedom within each species before and after binding states (eq. 5). In an induced fit helical interaction (3) such as this NR•SRC2-2 the peptide must change from a disordered peptide to a α-helix (eq. 6) with an entropic cost of $\Delta S_{peptide}$. This entropic cost can differ significantly depending on the composition of the peptide to alter the free energy $\Delta\Delta G_{int}$. We chose to quantify ΔS_{helix} by constraining the helix in advance (eq. 7) and consequently use the effective constraint method as a potential scaffold.



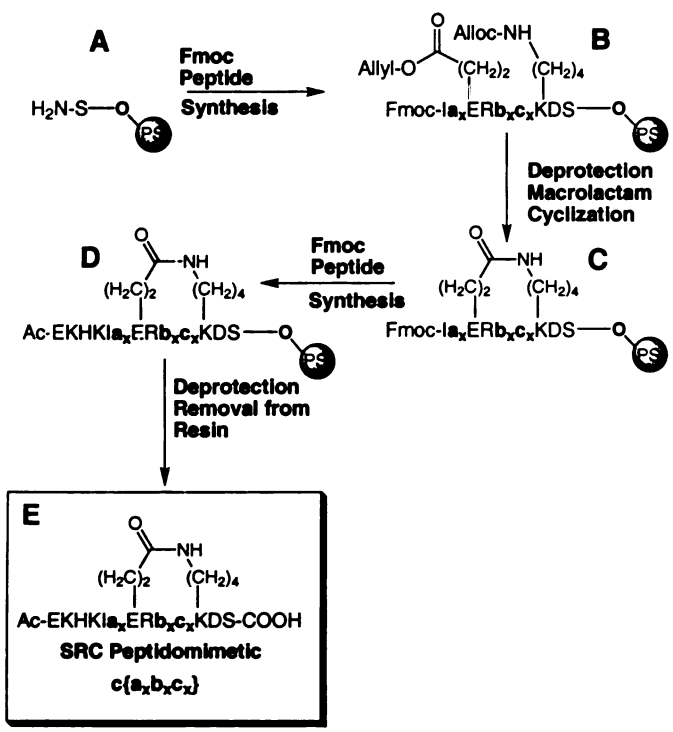
Synthesizing an Effective α-helical LXXLL Coactivatormimetic Library

In an induced fit α-helical interaction such as that proposed with NR•SRC one should constrain a peptide into a helix to remove the entropic cost of peptide helix formation ($\Delta S_{peptide}$) from the free energy of the binding interaction ΔG_{int}. Computational evaluation and some empirical evidence suggested that the ΔS of helix formation would be on the

order of 1 to 10 kcal/mole. However, it was not clear what this value would be in the presence of a protein interaction nor was it clear if this value would be consistent between different protein-protein interactions. For example, would ΔS_{helix} be similar for a given peptide sequence and its interaction between two highly homologous interfaces, hTRβ and hERα, for example. By constraining a sequence into an α-helix in advance and measuring its ability to compete with each protein relative to that of an unconstrained peptide, we should be able to measure the $\Delta\Delta G_{int\ (11vs.\ peptide)}$ as a function of a change in the $\Delta S_{peptide}$ with no changes in enthalpy. Applying the same principle to multiple different constraint methods one can identify an appropriate constraint to serve as a scaffold for the presentation of the leucine mimetic sidechains. Ideally this will remove the effect of sequence variation on helical propensity[35] and provide a more accurate evaluation of the contribution to binding by the direct interactions of the individual sidechains in the interface (Figure 2-1, eq.7).

Figure 2-2. Solid Phase Synthetic Scheme

Solid Phase Synthesis for the Lactam Constrained SRC2-2 Proteomimetic. Each synthesis involved the following steps: (A) Wang resin was loaded and Fmoc peptide synthesis (Addition of each additional amino acid followed a cycle of: 1) removal of the amino terminal Fmoc protecting group (20% piperidine / DMF / 22°C. 10min, two repetitions); 2) coupling of the next Fmoc protected amino acid to the growing peptide chain (HBTU (3eq) / DIEA (3-5eq) / DMF, 22°C, 0.5-2hrs, two repetitions) This proceeded until the peptide was two amino acids beyond the bridging residues to generate (B). Bridging residues were deprotected - orthogonal deprotection of the allyl and Alloc side chain protecting groups (Tetrakistriphenylphosphine palladium (0.0002eq) (STREM), N,N-dimethyl barbituric acid (NDMBA) (3eq), acetic acid (1.5eq), DCM, rt, 2hrs, three repetitions) - and the lactam constraint was formed - (HATU (3eq) / DIEA (3-5eq) / DMF, shaken, rt., 2hrs) - to produce (C). Fmoc peptide synthesis was continued until the end of the peptide and acetylated (D). Deprotection and removal from the resin under acidic conditions - cleavage from the resin and concomitant side chain deprotection (95%TFA, 1%H20, 2%Phenol, 2% Thioanisol, 22°C, 3hrs); 2) ether precipitation - generated the final SRC proteomimetic (E). Compounds were all purified to > 95% purity - HPLC Purification (RPC_{18} Column, 0.1% TFA in water to 0.1% TFA in 80% acetonitrile over 30 minutes). Structures were verified by MALDI HRMS to give satisfactory exact mass (see table below).



Based on the biochemistry and structural data, we chose the 13 amino acid segment of the SRC2-2 peptide as the background NH_2-^{685}EKHKI**L**$_1$ER**L**$_2$**L**$_3$KDS^{697}–COOH. The Kd of the peptide was known to be 1 μM [7,38,58] for hTRβ which was well within the measurable range for anisotropy binding and competition studies as well as GST pull down evaluations (see assays below). Many methods for constraining peptides into a α-helical conformation [27] were available and a few have been successfully applied to protein interfaces.[28-34]We chose to design a series of compounds with a constraint on the solvent exposed region of the peptide, which should not disrupt the hydrophobic face of the LXXLL motif. Each was evaluated for secondary structure by circular dichroism (See assays below) and for the ability to compete with the SRC2-2 binding interaction with T_3•hTRβ and E_2•hERα. One constraint was effective and was used as a scaffold for a second generation library that applied non-natural amino acids to each of the leucine positions sequentially to produce a coactivatormimetic {c(E^{691}-K^{695})Ac-^{685}EKHKI**a**$_x$ER**b**$_x$**c**$_x$KDS^{697}-COOH} with a macrolactam constraint at positions E^{691} and K^{695}, defined as (1{$a_xb_xc_x$}), that is presented in subsequent chapters three and four and shown to specifically inhibit the SRC2-2 interaction with several NR in both a NR and ligand dependent manner.

Results

A series of Coactivator-mimetics were synthesized with lactam bridges that have been shown to induce a α-helical conformation in some peptide sequences[31,32]. A series of macrolactam constrained GRIP 1 NR box peptides (Table 1) were synthesized by solid phase peptide synthesis (Figure 2-2) using the Fmoc strategy, with orthogonal protection of the relevant lactam precursor side chains, followed by on-resin formation of the macrolactam.[1] The compounds synthesized included variation in the location, length, and orientation of the lactam.

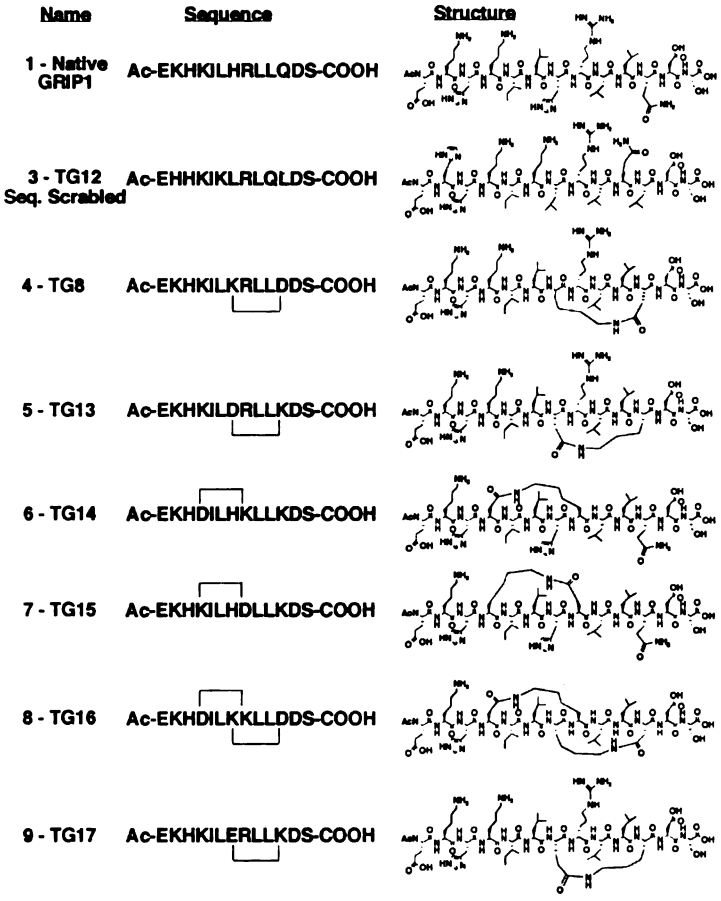


Figure 2-3. Scaffold Library

Several different lactam constraints were designed to stabilize the SRC2-2 (**1**) sequence into an α-helical conformation. Each constraint was designed around the core LXXLL motif and varied in position (**4** vs. **6**), direction (**4** vs. **5**), or length (**5** vs. **9**), for example. Each of the constraints were modeled and computationally evaluated for Molecular mechanics with MM3 force field minimization [NOTE:SYBYL® 6.7.1 Tripos Inc., 1699 South Hanley Rd., St. Louis, Missouri, 63144, USA] and measured for α-helical geometric violations.[84] Each passed without violating the geometry of the α-helix. A scrambled SRC2-2 sequence was generated for a negative control (**3**) and a fluorescent probe would be generated by labeling the SRC2-2 sequence with Oregon green at the N-terminus (**2**) (not shown) in place of the acetylation.

Solution phase circular dichroism (CD) analysis of these peptides (Figure 2-4) revealed partial induction of α-helical character in one, **9**, but little to no induction of helical character in the other compounds (Table 1). NOESY NMR experiments (Figure 2-5) with **9** revealed four amide proton resonances that show amide•amide (i, $i+1$) crosspeaks characteristic for peptide amides in an α-helical fold with COSY coupling constants less than 5.0Hz ($J_{HN-HC\alpha}$), well within the range normally exhibited in α-helical segments of proteins. These data imply either that peptide **9** has a helical conformation for 30-40% of its length or that populations of energetically similar conformations are interconverting rapidly. The structure of the CD spectra of unconstrained peptide **1** and constrained peptide **9** were independent of the temperature over the range of 4-50°C (Figure 2-4) supporting the former model. The combination of these findings suggests that the constraint in **9** strongly biases conformational equilibria towards an alpha helical structure within the portion of the peptide enveloped by the lactam.

Table 2-1. Sequence and observed Ellipticities of cyclic peptides.

Compound	$[\Theta]_{192}/[\Theta]_{222}$	$[\Theta]_{222}/[\Theta]_{208}$
Idea Helix	2.63	1.09
NH_2-^{685}EKHKILHRLLQDS^{697}-COOH **1**	-1.91	0.47
(Oregon Green 488)-EKHKILHRLLQDS-COOH **2**	NA	NA
Ac-EHHKIKLRLQLDS-COOH **3**	-4.09	0.37
c(D^{688}-K^{692},D^{691}-K^{695})Ac-EKHDILDKLLKDS-COOH **4**	-0.53	0.71
c(K^{691}-D^{695})Ac-EKHKILKRLLDDS-COOH **5**	-1.81	0.47
c(D^{691}-K^{695})Ac-EKHKILDRLLKDS-COOH **6**	0.20	0.47
c(D^{688}-K^{692})Ac-EKHDILHKLLQDS-COOH **7**	-0.53	0.71
c(K^{688}-D^{692})Ac-EKHKILHDLLQDS-COOH **8**	-0.20	0.81
c(E^{691}-K^{695})Ac-EKHKILERLLKDS-COOH **9**	1.41	0.70

The relative numbering scheme for GRIP1 amino acid residues is indicated on peptide **1**. The notation c(X-Y) implies formation of a macrolactam between the side-chains at the indicated positions. The NR box leucines are underlined. Linear probe and control peptides are above the bold line; constrained peptides are below the bold line.

Table 2-2. High Resolution Mass Spectroscopy Results

Compound	Calculated	Measured	Difference
1	1616.93	1617.00	0.07
3	1658.94	1658.98	0.05
4	1559.85	1559.95	0.10
5	1618.93	1619.96	1.03
6	1618.93	1620.02	1.09
7	1599.85	1599.90	0.05
8	1599.85	1599.83	-0.01
9	1632.95	1632.93	-0.02

Structures were verified by MALDI HRMS to give satisfactory exact mass (see table below). MALDI Instrument was a DE-STR-MALDI (Precision Biosystems, Framingham, MA). The matrix was MALDI-Quality Alpha-cyano-4-Hydroxycinnamic Acid in Methanol, (Hewlet Packard). Samples were prepared from 50uM peptide following the HPLC fractionation in ACN / H_2O in 0.1% TFA were mixed with the matrix at a 1:1 ratio.

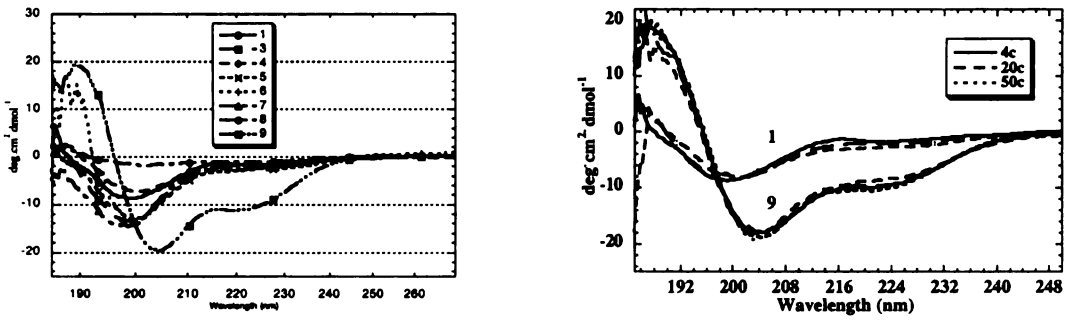


Figure 2-4. CD all at 4c and Temperature Dependence

All peptides were tested by CD. Circular Dichroism Spectra of all the linear and constrained macrolactam 13 amino acid peptides. This figure demonstrates that only 9 has any significant α-helix spectrum. Spectra were acquired by scanning solutions of each compound at 50μM in 20% ACN in 50mM Tris pH 8.0. Mean residue ellipticity (Θ) reported in deg·cm^{2}·dmol^{-1}. Temperature Dependence and circular dichroism spectra of the unconstrained peptide 1 and constrained peptide 9. Spectra were acquired by scanning solutions of 1 and 9 (50μM in 20% ACN in 50mM Tris pH 8.0) at various temperatures. Mean residue ellipticity (Θ) reported in deg·cm^{2}·dmol^{-1}.

The ability of the constrained peptides to displace native GRIP1 peptide from E_2•hERα and T_3•hTRβ was assessed using fluorescence polarization (FP) equilibrium competition assays (Figure 2-6). Control experiments indicated that binding of probe **2** was dependent upon the presence of ligand E_2 or T_3 (♦) and that competition for binding to hERα and hTRβ was exhibited by unconstrained peptide **1** (▼) but not by control peptide **3,** which has a sequence scrambled NR box (■). The ability of each of the constrained GRIP 1 analogs to successfully compete for binding was not consistent between the two NRs.

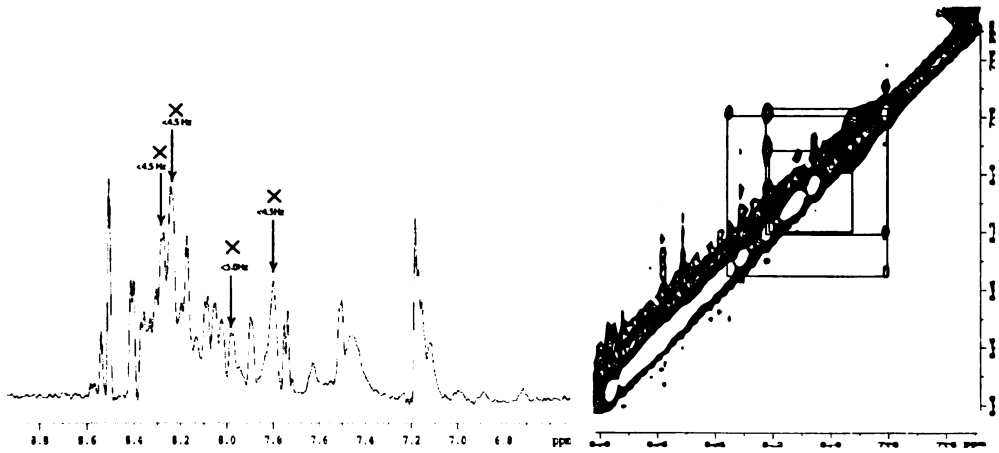


Figure 2-5. COSY Spectra Peptide 9 - COSY 1D ^{1}H-NMR with ($J_{HN-HC\alpha}$) Coupling Constants and NOESY 2D ^{1}HNMR (N_iH to $N_{i+1}H$) Peptide 9

Measured Coupling constants for the 4 HN to HC_α - COSY coupling constants of <5.0Hz ($J_{HN-HC\alpha}$) clarified that these protons identified in the NOESY spectra are consistently in a cis orientation to each other enforced by the secondary α-helical conformation for 9.

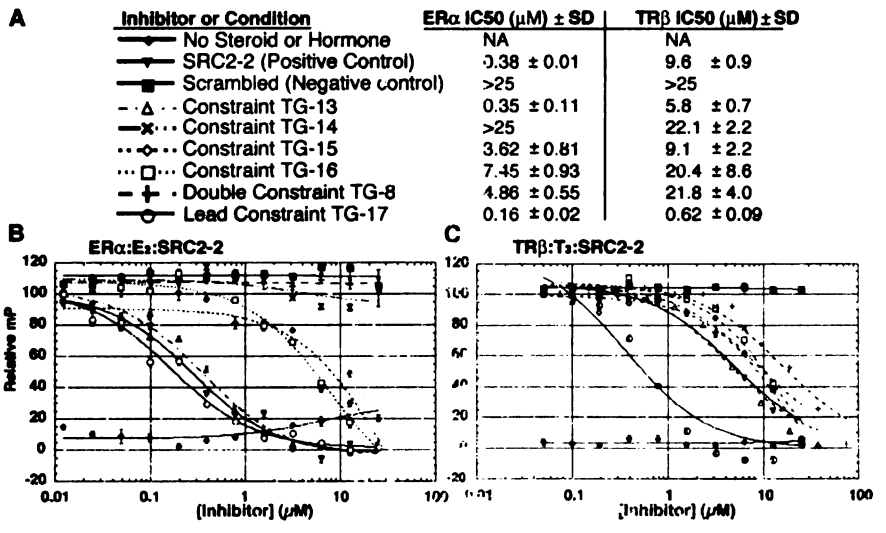
Not all of the constrained compounds were able to compete with GRIP1 NR box 2 (SRC2-2) for binding to hERα and none were significantly better than the linear SRC2-2. In contrast, with hTRβ, all of the constrained analogues were able to compete effectively with GRIP1 NR box 2 (SRC2-2) at the hTRβ interface and, strikingly, **9**, (○) the single constrained peptide that exhibited substantial helical character, exhibited an IC_{50} 15-fold lower than the unconstrained peptide **1** (▼). The change in inhibitory constant exhibited by **9** indicates that formation of the helix during binding induces a 1.5 kcal/mole cost in free energy of binding to hTRβ. The magnitude of this effect is equivalent with that seen in studies of the yeast transcription factor GCN4 when similar constraints were applied.[2] This same compound also successfully competed for binding to hERα, however, with only a two fold improvement (approximately 0.44 kcal/mol) over that of the non-constrained SRC2-2 peptide. This study demonstrates that constraint of the NR box to an α-helical structure strongly enhances affinity of the NR box for the hTRβ receptor and only mildly with that of the hERα receptor. This observed difference may be due to steric clash between the lactam constraint and hERα but not with hTRβ, or the final constrained structure may not be energetically ideal for hERα.

The ability of constrained peptide **9** to compete with intact GRIP1 nuclear receptor interaction domain (GRIP1 NID), which contains all three GRIP1 NR boxes, was tested using a semi-quantitative glutathione-S-transferase assay (Figure 2-7). Control experiments indicated that the GRIP1 NID bound to hTRβ in the presence of T_3 (lane 3) and failed to bind in the absence of T_3 (lane 2). This interaction was blocked by unconstrained peptide **1** at high concentration (lane 5) but not at low concentration (lane 4) and was not blocked by control peptide **3** (lane 6). Constrained peptide **9** efficiently blocked the binding of hTRβ to the GRIP1 NID in a dose dependent manner (lanes 7-10). Although the assay does not allow for exact determination of IC_{50} values, the relative potency of **9** to **1** was qualitatively in the same range as that observed in the FP studies. Additionally, the interaction was completely blocked by 10 μM **9** whereas the unconstrained peptide never reached this level of saturation, even with 100 μM

concentrations. This study indicates that the constrained peptide **9** can block the interaction of intact receptor and coactivator proteins.

Figure 2-6. Fluorescence Polarization Competition Assay TR and ER

Inhibition of SRC2-2 binding to hTRβ and hERα by constrained SRC2-2 analogues. (A) Six constraints (TG-8, 13, 14, 15, 16, 17) were evaluated under similar conditions with two nuclear receptors ERα (B) and TRβ (C) for their ability to compete with native SRC2-2 peptide. Assay determined by fluorescence polarization competition experiments using an Oregon Green 488 labeled SRC2-2 (FITC-SRC2-2) peptide as previously described [81]. (◆) SRC2-2 probe does not bind in the absence of ligand E_2 or T_3 with ERα and TRβ respectively; (▼) unconstrained **SRC2-2** linear peptide specifically blocks binding of the FITC-SRC2-2 probe with both receptors hERα-IC50 0.38±0.01μM and hTRβ-IC50 9.6±0.9μM; (■) **Scrambled SRC2-2** control peptide is unable to block binding of FITC-SRC2-2 with either receptor; (+) (**Constraint TG-8**) hERα-IC50 4.86±0.55μM and hTRβ-IC50 2[illegible].8±4.0μM; (Δ) **Constraint TG-13** hERα-IC50 0.35±0.11μM and hTRβ-IC50 5.8±0.7μM; (X) **Constraint TG-14** hERα-IC50 >25.0μM and hTRβ-IC50 22.1±2.2μM; (◊) **Constraint TG-15** hERα-IC50 3.62±0.81μM and hTRβ- IC50 9.1±2.2μM; (□) **Constraint TG-16** hERα-IC50 7.45±0.93μM and hTRβ-IC50 20.4±8.6μM; (○) **Constraint TG-17** hERα-IC50 0.16±0.02μM and hTRβ-IC50 of 0.62±0.09μM. FITC-SRC2-2 probe was held constant at 10nM, [hTRβ] 1.5μM, [T_3] 10μM or [illegible] 10μM. Binding buffer: 20mM Tris HCl pH 8.0, 100mM NaCl, 10% glycerol, 1mM DTT, 0.01% NP-40, 1mM EDTA.



Discussion and Conclusions

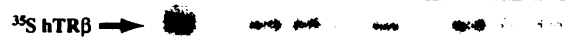
Our results suggested that the formation of a Glu to Lys macrolactam $c(E^{691}\text{-}K^{695})$ in the second NR box of GRIP1 induces a partial α-helical conformation. This conformational constraint allows **9** to compete for the NR box binding site of hTRβ with a 15-fold decrease in IC_{50} and hERα with a 2-fold decrease in IC_{50} relative to GRIP1 thus demonstrating that pre-forming an α-helical conformation in the NR-box leads to stronger interaction between the coactivator and the receptor. The poor competitive abilities of other constrained NR box peptides are most likely due to the lack of α-helicity as indicated by solution CD, thus confirming the requirement for an α-helical conformation in the NR box during binding to the nuclear receptor. The ability to block formation of the T_3•hTRβ•GRIP1 complex should allow functional antagonism of thyroid

hormone signaling. Thus, the results of this study suggest that this interface is an appropriate target for the development of protein-protein inhibitors as novel thyroid hormone receptor antagonists.

Figure 2-7. GST Pull Down Competition Assay

Inhibition of GRIP1 NID protein binding to hTRβ by **9.** Lane: (1) ^{35}S-hTRβ alone; (2) No T3 hormone, ^{35}S-hTRβ binding is ligand dependent; (3) no competitor maximal binding of ^{35}S-hTRβ to GRIP 1 NID domain; (4) 0.01 μM **1**; (5) 100μM peptide **1** will compete for binding to hTRβ; (6) 100μM **3** showed no competition for binding to hTRβ; (7 - 10) increasing concentration of **9** conformational constraint increases competitive ability for binding to ^{35}S-hTRβ. ^{35}S-hTRβ was *in vitro* expressed with ^{35}S-methionine. Recombinant fusion of glutathione-S-transferase and hGRIP1 (GST-GRIP1(563-767-H6)) bound to glutathione agarose beads was exposed to ^{35}S-met-hTRβ in the presence and absence of inhibitor. Binding buffer was 20mM Tris HCl pH 8.0, 100mM NaCl, 10% glycerol, 1mM DTT, 0.01% NP-40, 1mM EDTA.

Lane	1	2	3	4	5	6		7	8	9	10
^{35}S-hTRβ	+	+	+	+	+	+		+	+	+	+
GST-GRIP1		+	+	+	+	+		+	+	+	+
T3			+	+	+	+		+	+	+	+
1 (μM)				0.01	100						
3 (μM)						100					
9 (μM)								0.01	0.1	1.0	10.0



Why we see such a difference between the interfaces of hERα and hTRβ is not entirely clear in the absence of structural data. One could expect that the relative response to the library of constrained compounds would have been similar given the degree of conservation between the two interfaces. The lead compound TG-17 **9**, for example, provides only a 2-fold improvement in the IC_{50} with hERα compared to the 15 fold seen with hTRβ. One possibility for this difference is that this constrained SRC2-2 compound interacts with the two receptor interfaces differently and that the lactam constraint is creating a slight perturbation to the presentation of the leucines or is sterically hindering the interaction with hERα. Another possibility is that there is a difference in the relative benefit of the preformed, solution phase, α-helicity for binding to each receptor. The differences seen between the two receptors may be due to the ability one interface to engender the α-helicity of this NR box sequence differently from the other receptor interface. Consequently this may suggest that it is the ability of the NR to engender helicity of a given NR box sequence differently from that of another NR and regulate binding. The sequences flanking the NR box may form transient interactions with each receptor surface differently to seed helicity. hERα may do this more effectively than hTRβ with the SRC2-2 sequence.

Additionally we see that several of the constrained compounds are able to compete with the interface of the receptors quite effectively despite not being helical in solution.

Unfortunately, the binding mode of these interaction is unclear in the absence of structure. Presumably these compounds are constrained into a conformation that is non-helical in solution, as determined by CD, but are able to adopt an α-helix upon binding to the hERα interface without paying too much of a penalty relative to the linear compound or even the lead compound. These compounds may be stabilized into a 3_{10}-helix, for example, which would not be detectable by CD in solution but could possibly make the transition to an α-helix. The formation of a 3_{10}-helix has been hypothesized to be an important low energy transition state between the non-helical and α-helical state of peptides. This may be particularly true at the interface of protein-protein interactions involving such an induced fit α-helix. Based on these last two considerations we see again that it may be possible that NR box helicity is playing a significant regulatory role in determining the relative on and off rates between each NR coregulator NR box interaction and may identify the role of the sequences flanking the LXXLL motif which seem to determine selectivity between the receptors [8,77].

Methods

Synthesis of Peptides and Constrained Proteomimetics

Orthogonally protected peptides **1** and **3-9** were synthesized on solid support from the carboxy to amino terminus with the carboxyl group of the first amino acid attached to Merrifield resin. Each synthesis involved the following steps: Addition of each additional amino acid followed a cycle of: 1) removal of the amino terminal Fmoc protecting group (20% piperidine / DMF / 22°C. 10min, two repetitions); 2) coupling of the next Fmoc protected amino acid to the growing peptide chain (HBTU (3eq) / DIEA (3-5eq) / DMF, 22°C, 0.5-2hrs, two repetitions). Coupling efficiency during the peptide synthesis was monitored by the Kaiser test and coupling was continued until the resin failed to show a positive blue color. Formation of the lactam bridging group followed the sequence of: 1) orthogonal deprotection of the allyl and Alloc side chain protecting groups (Tetrakistriphenylphosphine palladium (0.0002eq) (STREM), N,N-dimethyl barbituric acid (NDMBA) (3eq), acetic acid (1.5eq), DCM, rt, 2hrs, three repetitions); 2) lactam formation (HATU (3eq) / DIEA (3-5eq) / DMF, shaken, rt., 2hrs). Coupling

efficiency was monitored via the Kaiser test until the resin failed to show a positive blue color. Amino acids were purchased from Nova Biochem, San Diego, CA. Allyl protected aspartic acid and glutamic acid were synthesized separately (See below). Final steps followed the sequence of: 1) cleavage from the resin and concomitant side chain deprotection (95%TFA, 1%H20, 2%Phenol, 2% Thioanisol, 22°C, 3hrs); 2) ether precipitation; 3) HPLC Purification (RPC_{18} Column, 0.1% TFA in water to 0.1% TFA in 80% acetonitrile over 30 minutes). Structures were verified by MALDI HRMS to give satisfactory exact mass (see table below). MALDI Instrument was a DE-STR-MALDI (Precision Biosystems, Framingham, MA). The matrix was MALDI-Quality Alpha-cyano-4-Hydroxycinnamic Acid in Methanol, (Hewlet Packard). Samples were prepared from 50uM peptide following the HPLC fractionation in ACN / H_2O in 0.1% TFA were mixed with the matrix at a 1:1 ratio.

Preparation of Allyl Protected Amino Acids Fmoc-Asp(Allyl)-OH and Fmoc-Glu(Allyl)-OH

Fmoc-Asp-O-tBu and Fmoc-Glu-tBu amino acid sidechains were protected by an allyl group in solution phase chemistry. The t-butyl group was subsequently removed to provide an Fmoc-Asp(Allyl)-OH and Fmoc-Glu(Allyl)-OH aminoacids for peptide synthesis. Activation of the acid side chains utilized DMAP and DCC with attack of the allyl alcohol. To a dichloromethane (DCM) solution (100ml) 4g of the respective amino acid, 1eq. dicyclohexylcarbodiimide (DCC), 0.05eq. dmethylaminopyridine (DMAP) and 5 eq. of the allyl alcohol were added and stirred for 2 hours. Allyl protected aminoacids were extracted in 1M NH_4Cl (pH 4) and purified by silica gel chromatography (10% ethyl acetate/hexane followed by 20%EA/Hexane) for a yield of 68% as a yellow oil. An equal volume of TFA (40 ml) was added to the oil and stirred for 2 hours. 160 ml of Toluene was added and stirred for 5 minutes, and roto-evaporated off. Toluene addition and roto-evaporation was repeated three times to finally yield a white powder at 85% yield. Analysis: ^1H NMR (500MHz, $CDCl_3$) Fmoc-Asp(Allyl)-OH 100mg/ml of the final product Fmoc-Asp(Allyl)-OH in $CDCl_3$ was confirmed in a 1D ^1H NMR (500MHz) resulting in the following peaks: 3.01 (q 2H), 4.46 (t 1H), 4.70 (d of t 2H), 4.75 (d 2H), 5.06 (t 1H), 5.24 (d of d 2H), 5.81 (d 1H), 5.90 (octet 1H) 7.29 (t 2H),

7.38 (t 2H), 7.58(d 2H), 7.75 (d 2H). ^{1}H NMR (500MHz, $CDCl_3$) Fmoc-Glu(Allyl)-OH. 100mg/ml of the final product Fmoc-Glu(Allyl)-OH was dissolved in $CDCl_3$ and tested in a 1D ^{1}H NMR (500MHz) resulting in the following peaks: 1.95 (q 2H) 2.25 (q 2H), 4.46 (t 1H), 4.7 (d of t 2H), 4.75 (d 2H), 5.06 (t 1H), 5.24 (d of d 2H), 5.81 (d 1H), 5.90 (octet 1H) 7.29 (t 2H), 7.38 (t 2H), 7.58(d 2H), 7.75 (d 2H)

Synthesis of Labeled Probe 2

A solution of the peptide **1** (0.8 μmoles) in aqueous DMF (100 μL 1:1, 200 mM sodium phosphate (pH 7): DMF) was treated with Oregon Green 488 succinimidyl ester (Molecular probes 0-6147, 3.2μM in 50μL DMF) and allowed to stir at room temperature for 2 h. The resulting fluorescently labeled peptide **2** was purified by RP-HPLC as above.

Circular Dichroism

Circular dichroism spectra were collected on a Jasco spectropolarimeter. A solution of peptide (50 μM in 20% acetonitrile in 50 mM Tris•HCL, pH 8.0) was placed in a 1 cm pathlength cuvette and its circular dichroism specta measured. The spectra was normalized using a solution of buffer without peptide. Temperature studies utilized a circulating water bath temperature control and direct measurement throughout the data collection.

H6-hTRβ-LBD Expression and Purification

hTRβ LBD was expressed in *Escherichia coli* DH5α. Frozen competent DH5α Cells were thawed and transformed with a PET3-TRB-LBD in $CaCl_2$. Cells were grown up at 14° C, 1 mM IPTG added at OD600 = 0.7, induced 24 hr at 14°). Cells were lysed by freeze-thaw in 50 mM sodium phosphate (pH 8.0), 0.3 M NaCl, 10% glycerol, 25 mM β-mercaptoethanol, and 0.1 mM PMSF after incubation with 0.1 mg/ml lysozyme (20 min, 0° C). The lysate was cleared (Ti45, 36,000 rpm, 1 hr, 4° C) and the protein purified using the procedure described above, except in sodium phosphate buffer, and eluted with 12-300 mM imidazole gradient. The liganded [(3,3',5-triiodo-L-thyronine (T3; Sigma)] receptor was isolated by HPLC (TSK-phenyl column (TosoHaas, Philadelphia, PA)).

Imidazole was removed by NAP gel filtration (Pharmacia) and protein concentration determined by BCA assay (Pierce).

hERα Protein expression and purification

The human ERα LBD (residues 297-554) was expressed in BL21(DE3), carboxymethylated and purified on an estradiol-sepharose column as described*. Protein was eluted with E_2 30 μM in 30 to 100 ml of 50mM Tris, 1 mM EDTA, 4 mM DTT, 250 mM NaSCN , 10% N,N-DMF (pH 8.5). The buffer was exchanged into 20mM Tris, 1 mM EDTA, 4 mM DTT (pH 8.1) to remove denaturants, concentrated to ~ 1mg/ml with ~10 μM free ligand in the buffer and stored at –20c till further use. Protein was analyzed by SDS-PAGE, size exclusion HPLC (SEC 125-5, Bio-Rad), and MALDI-tof Mass spectrometry (Voyager-DE, PerSeptive Biosystems). Procedure was adapted from Shiau, A.K. et al. Cell 95, 927-937 (1998).

Fluorescence Polarization Assay of Competitive Ability

All spectral data were acquired using an LJL Biosystems Analyst AD system. Pilot experiments demonstrated that the binding of **2** to TR was completely saturable and reached equilibrium within 10 minutes. Our observed K_D agreed well with the reported value of 1 μM. Preliminary competition revealed that **1** could completely compete with the labeled probe for binding to TR and that this equilibrium was also reached within 10 minutes. Because anisotropy inherently measures a ratio of bound to unbound probe, competition experiments must be carried out at high receptor concentrations, ideally two fold that of the K_D.

These studies established the validity of several controls: a positive control (competition by unlabeled GRIP1 NR Box peptide **1**), and negative controls for hormone dependence of probe binding (no thyroid hormone) and for unsuccessful competition (a sequence scrambled NR box peptide that breaks up the leucine triad, peptide **3**). We have also carefully studied the effects of pH and salt concentration upon the system and settled upon an optimal buffer (20 mM Tris pH 8; 100 mM NaCl, 1 mM DTT, 1 mM EDTA, 0.01% NP-40, protease inhibitors).

Experimentally, potential competitors are evaluated by pre-equilibrating the receptor, probe, and hormone for one hour, titrating a set of samples with increasing concentrations of each inhibitor, and then evaluating for displacement of the probe after one hour. All experiments are carried out in quadruplicate, with each iteration containing the positive and negative competition controls and twelve dose points. The data are then fit using Klotz plots to determine IC_{50} values using nonlinear regression analysis that fits the data to a modification of the model of Heyduk and Lee.(Heyduk and Lee 1990; Heyduk, Lee et al. 1993; Heyduk, Ma et al. 1996)

GST Pulldown Assay Qualitative Assessment of Competitive Ability

GST binding assay was described previously (Hong, H.; Darimont, B. D.; Ma, H.; Yang, L.; Yamamoto, K. R.; Stallcup, M. R. *J. Biol. Chem.* **1999**, *274*, 3496-3502.). The GRIP1 GST-NID derivative was expressed in BL21(DE3) cells. Cells were thawed and $CaCl_2$transformed with the PGEX-$GRIP_{1565\text{-}767\text{-}H6}$ plasmid. Cells were grown up at 37° C, 1 mM IPTG was added at OD600 = 0.7 to induced for 4 hrs. Cells were lysed by sonication in sonication buffer (20 mM Tris-HCl,pH 8.0, 0.1 M NaCl, 10% glycerol, 0.1 mM PMSF, protease inhibitors, (Complete, EDTA free, Boehringer Mannheim). The protein was purified by affinity chromatography on Talon resin (Clontech) equilibrated in sonication buffer. Crude protein was loaded onto the column at 1ml/min, washed with the sonication buffer plus 4mM Imidazole until chromatograph returned to baseline, washed again with 12mM imidazole sonication buffer solution and eluted with a linear gradient of imidazole from 12to100 mM over 100min yielding 95% pure protein. Protein was analyzed by SDS-PAGE with both comassie and silver staining. Imidazole was removed by NAP gel filtration (Pharmacia) and protein concentration determined by BCA assay (Pierce).

The GST-NID was bound to glutathione•agarose (Sigma) in binding buffer (1 hr, 4° C, sonication buffer + 1 mM DTT, 1 mM EDTA, 0.01% NP-40). Excess protein and impurities were removed by washing with 4x5 volumes of binding buffer. The beads were then diluted in binding buffer with 20% glycerol to 40% final bead concentration, snap frozen in liquid nitrogen, and stored at –70° C. Loaded beads were analyzed by

SDS-PAGE using BSA standard (Pierce) to ensure equal concentrations of bound proteins.

The PET3-TRB-LBD hTRβ plasmid was used to express ^{35}S-labeled hTRβ in a coupled transcription-translation rabbit reticulocyte lysate system (TNT; Promega). Translations were performed in the presence and absence of 10 μM T3. Expression of hTRβ yielded 20-100 ng/μl reaction, as computed from the amount of incorporated ^{35}S-methionine.

Fifty microliters of 20% bead suspension containing 4.0 μM GST-NID was incubated with 0.2-2 μl of lysate containing labeled ^{35}S-hTRβ ; final concentration of receptors was 10 nM. Binding (4° C under rotation, 2 hr) was in binding buffer + 20 μg/ml BSA, with or without 10 μM T3. For competition experiments, peptides were added before receptors, but altering the order of addition did not affect the results (data not shown), confirming that the reactions reach equilibrium. Beads were washed (4° C, 5 x 200 μl binding buffer), proteins eluted (10 μl 2x SDS loading buffer), subjected to SDS-PAGE, and the fraction of bound receptor determined (Molecular Dynamics Phosphor-Imager, Storm 860). The receptor binding was independent of lysate concentration.

NMR Spectroscopy Peptide 9

Peptide 9 was lyophilized down from the HPLC elution (ACN / H_2O in 0.1% TFA) and re-suspended in the buffer and D_2O and placed in a Shigemi NMR tube. (1mM peptide 9 in 300ul of 50mM Phosphate buffer, pH 5.0, 10% D_2O). The instrument utilized was a Varian (600MHz), with an Oxford Magnet, Unity and Inova Hardware. Spectra were collected at 277K and 298K with no significant differences.

Acknowledgments

The authors thank R. Fletterick and K. Yamamoto for informative discussions, B. Darimont for technical assistance, V. Doetsch for acquisition of the NMR data, and the Sidney Kimmel Foundation for Cancer Research, the HHMI Research Resources Program grant #76296-549901 and the Academic Senate of the University of California at San Francisco for financial support.

Chapter III Computational Analysis of the Nuclear Receptor : Coactivator Interface - Identifying Sub-site differences between ERα, ERβ and TRβ for inhibitor design

Timothy R. Geistlinger.; R. Kiplin Guy, Steroid Receptor Coactivator Proteomimetics, *Methods In Enzymology*, **2003**, Vol. 364, Chapter 13, pp 223-246

Timothy R. Geistlinger.; Andrea McReynolds,; R. Kiplin Guy, Specific Inhibitors Select for Different Ligands with Each Estrogen Receptor Isoform, *Chemistry and Biology*; **2004**; (11) pp 273-281.

Introduction

Several fundamental properties of protein-protein interactions have hindered development of drugs directed at these targets. These hindrances include a paucity of lead compounds, large interface sizes that seem to demand "large molecules", high degrees of structural plasticity in the targets that hinder rational drug design, and topological simplicity of the often highly conserved interfaces that leave medicinal chemists few handles for building selective drugs. Due to the relatively small volume and rigid character generally desired in drugs, the latter two characteristics conspire to make the production of receptor isoform selective compounds particularly difficult. Therefore, the determination of whether or not a particular interface allows for the development of selective inhibitors is paramount in early drug discovery studies for these targets.

While each of the leucines in the NR box is critical to the interaction,[61] specificity in recruiting a particular SRC appears to be induced by the sequences immediately flanking the NR boxes[8,64,75,76] rather than the geometry of the leucine side chains. Manipulation of NR box peptide sequence outside of the conserved $L_1XXL_2L_3$ motif has afforded selective peptide inhibitors of the interaction of particular NR and SRC's, presumably by taking advantage of the extended interactions naturally used to induce selectivity *in vivo*.[8,77,78] However, such extended elements would not be available to a small molecule disrupting the interaction and it is therefore uncertain how such findings reflect upon the development of such a drug. As a result, it is paramount to identify sub-site specific differences within this small hydrophobic pocket which could be effectively translated into a small molecule. The identification of selectivity within this relatively small motif, unlike other methods that flank this region could be effectively mimicked with a small molecule inhibitor. The crux, however, is identifying significant differences within this relatively small, highly conserved, hydrophobic pocket which act as handles for the small molecule. We therefore utilized the atomic resolution crystal structures of different NR for computational evaluation to first identify subtle differences in the interface of each NR. Then using *in silico* combinatorial docking (CombiDOCK) we designed a library of

non-natural leucine mimetic side chains. The hypothesis was that non-natural sidechains could take advantage of subtle differences that the native 20 amino acids could not select between. Through computational measurement we could identify differences that were subtle, seemingly insignificant, and more efficiently design a library of molecules to exploit these differences. We therefore rationally designed a series of SRC2-2 NR box LXXLL mimetic inhibitors [79-83] with one of three leucines substituted on an α-helical constrained SRC2-2 proteomimetic scaffold.[81]

Analysis of The Coactivator Binding Pocket of NR

Phylogenic and computational analysis of different NR hydrophobic SRC recognition pockets demonstrates that there are significant differences in amino acid composition of the residues making up the hydrophobic pockets on each receptor which create differing electrostatic potentials and topologies throughout the pocket (Figure 2). A total of 19 amino acid side chains on the surface of both hERα and hTRβ are either buried by SRC2-2 or form direct contacts with the NR box (Figure 2A), however, they are not always in the same respective location or of the same composition. Only eight of these amino acid residues are sequentially conserved between the two receptors. Fifteen of the 19 residues are different between TRβ and ERα by sequence or degree of physical interaction with the SRC2-2 NR box leucines, (L354I; V355T; M357V; I358V; V364L; V368C; L372C; V376I; E380K; D538P; M543V F367F and Q375Q - ERα numbering see Figure XX), Eight of these form unique direct contacts (underlined) with the leucines of SRC2-2 in one receptor and not the other, while two residues (C381G and A382C) are sequentially different and buried in the TR pocket but not with ERα. Four of these (L354I M357V I358V V364L) are conservative substitution while six (V355T; V368C; L372C; C381G; A382C; D538P) are positions with large changes in shape, electrostatics, or hydrogen bonding potential. The presence of a large number of surface exposed cysteines on TR in the coactivator binding pocket is somewhat surprising and may point to a potential for regulation of coactivator binding by post-translational modifications. The relative location and substitutions of P and G between ER and TR, are of some interest and may reflect potential for greater flexibility of secondary structure in ER. In general, the large number of significant changes in residues involved in coactivator

binding suggested that achieving selectivity between binding to liganded TR and ER would be achievable.

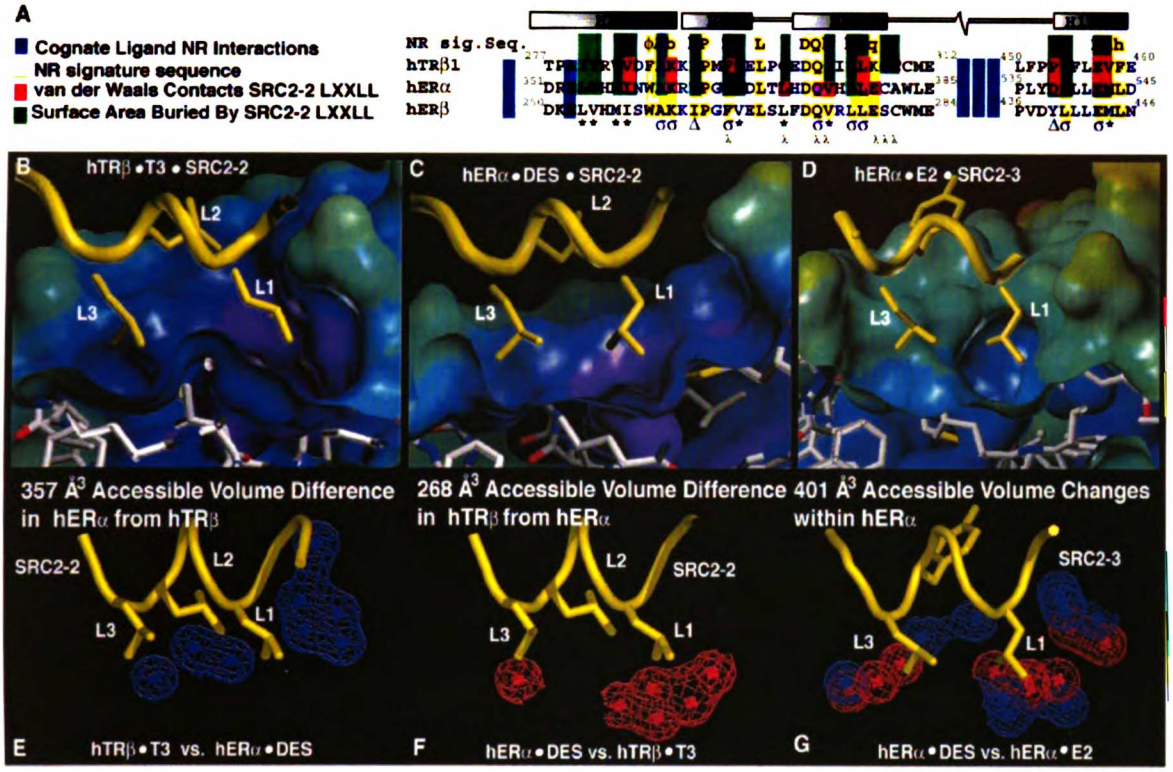


Figure 3-1. Phylogenic and Structural Comparison

Phylogenic and structural comparison of three co-crystal structures highlighting significant differences in the sub-sites of each nuclear receptor interface (Panel A) Alignment of the primary amino acid sequence that make-up the coactivator binding pockets of the receptors hTRβ, hERα and hERβ. The secondary structure of the proteins is identified in gray with each helix numbered according to the convention for NR. Amino acids are shown in single letter format with residue numbers from the receptor primers beginning and ending each segment. The NR signature sequence is shown on top with (ϕ) indicating an aromatic residue, (A) indicates alanine, (b) indicating a basic residue, (h) indicating any hydrophobic residue, (P) indicating proline, (L) indicating leucine, (Q) indicating glutamine, and (q) indicating a charged residue. Blue denotes areas of the receptor that make contact with the respective native ligands of each receptor. Red denotes residues identified to form VDW interactions with the leucine side chains of SRC2-2 in the respective crystal structures. Green denotes residues that compose the hydrophobic pocket and are buried in the crystal structures. (*) indicates residues that are buried in the respective complexes and are different between the hTRβ and the two ER isoforms but are the same between the isoforms. (σ) Indicates the residues that are the same across all the receptors, and (Δ) indicates residues that different across each receptor. This clearly demonstrates differences in the side chain character of the SRC binding pocket. (λ) Residues that have a different type of interaction with the coactivator. (Panels B-G) Co-crystal structure of the SRC NR box binding pocket in: (B) T_3•hTRβ•SRC2-2[7] (PDB:1BSX); (C) DES•hERα•SRC2-2[18] (PDB:3ERB); and (D) E_2•hERα•SRC2-3[66] (PDB:1GWR). The SRC binding pocket in each co-crystal structure is shown as a cross-section through the L_1 and L_3 recognition pockets to highlight structural and electrostatic differences within the highly conserved binding pockets. The three leucine side chains of the SRC2-2 ($L_1XXL_2L_3$) and the SRC2-3 (LLTXXLL) motifs (yellow wire frame) and labeled L_1 L_2 L_3. Each receptor Conolly surface was generated with a probe of 1.4Å radius and color coded to induce a continuous electrostatic gradient (*Sybyl*™). These views demonstrate clear significant differences in steric structure and electrostatics between the hTRβ and hERα SRC binding pockets (Panels B vs. C) as well as potential plasticity within hERα surface (Panels

C vs. D). These structural and volumetric differences between each comparison are highlighted in panels E and F with difference maps consisting of 2.2Å radius spheres that differentially fill each pocket of each receptor.[85] (Panels E - G) Oriented the same as panel B – D (*PyMOL*) spheres of 2.2Å radius filled the solvent accessible volume of each receptor pocket. Spheres were compared and removed if shared between receptor pockets leaving only those that were present in one structure and not the other. Panel E) Blue web spheres highlight the location of eight spheres totaling 357 $Å^3$ of volume that is available in the hERα pocket and not in the hTRβ pocket. Panel F) Six red spheres locating 268 $Å^3$ of volume that is available in hTRβ and not hERα. Panel G) A similar comparison between the pockets of (C) DES•hERα•SRC2-2 and (D) E_2•hERα•SRC2-3 with blue denoting the location of nine spheres totaling 401 $Å^3$ present in the pocket of DES•hERα•SRC2-2, and not in E_2•hERα•SRC2-3; and the reverse comparison also showing nine red spheres but in different locations. This highlights that the pocket has changed in topography without a change in accessible volume, possibly due to the change in ligand or in response to the different SRC interaction motif.

The two ER isoforms have 64% overall homology with 59% homology in the LBD[86] suggesting that achieving selectivity between them within this interface would be more difficult than with TRβ. Assuming that ERα and ERβ interact with the SRC2-2 peptide similarly – no co-crystal structure of ERβ•SRC2-2 is available - only two of the 19 residues lining the pocket are different between the ER isoforms. V364I is a fairly conservative change while D538Y (hERα numbering) is a major change. The Cα position of both of these residues deviate between hERα and hERβ by an RMSD of <0.2 Å. Together, this suggests that the sidechains are uniquely defining the pockets of each receptor with no compensation by backbone changes. Two substitutions (C381S and A382C) present changes in electrostatic or hydrogen bonding potential. These findings imply that achieving selectivity in binding to the coactivator binding pockets of the ER isoforms would be possible yet difficult.

Structural Analysis

Structurally, the interactions of hERα and hTRβ with the SRC2-2 peptide have been shown to be very similar to that of other NR•Coactivator peptide crystal structures solved to date (Figure 3-1B,C).[7,18,67,72,73,75] Each seem to interact through similar surfaces with a shallow hydrophobic groove on the NR surface binding to an induced fit, amphipathic, α-helical motif on the SRC2-2 NR box, burying the three leucines on the hydrophobic face of the NR box helix. At the protein fold level, the differences between the NR coactivator interfaces (Figure 3-1 B-D) of T_3•hTRβ•SRC2-2,[7] DES•hERα•SRC2-2,[18] and E_2•hERα•SRC2-3[66] (PDB:1BSX, PDB:3ERB, and PDB:1GWR). appear minimal. Each of the SRC2 peptide main chain RMSDs are less than 1.0 Å and the receptor main chain RMSDs are less than 2.0Å. However, upon closer investigation we find that there are

differences in the side-chain positions within each pocket which alter the electrostatics, topography and the location of solvent accessible volumetric differences between the pockets of each receptor (Figure 3-1E-F). In particular, a comparison between hTRβ and hERα we see hERα (Figure 3-1C) has a ridge with a continuous electrostatic potential adjacent to the L_2 sub-pocket whereas hTRβ (Figure 3-1B) has a sub-pocket that is sterically hindered by a more significant plateau with a gradient of electrostatic potential. Additionally, a pronounced ridge in hTRβ creates more steric hindrance between the L_1 and L_3 sub-pockets than what is present in hERα. This difference is due to a slight twist in helix 5 of TRβ and a substitution in the amino acid I302 of TR to a V376 in ER. The volumetric differences between these two receptors were determined to be 357Å^3 available in the ERα pocket that was different from the hTRβ pocket (Figure 3-1E) and 268Å^3 available in hTRβ and not in hERα (Figure 3-1F). These differences should be significant enough to computationally identify non-natural residues that differentially exploit these structural and volumetric differences.

The nuclear receptor has a distinct biological response to different natural and synthetic estrogen-like ligands. [14] [15-17] Each ligand seems to have a different effect on the estrogen receptor isoforms. But are these differences more subtle than the relative position of helix 12, and could we identify subtle differences in the coactivator binding pocket between different agonist ligands? Generating inhibitors to the protein-protein interaction between the SRC and each ER isoform that were selective not only to each isoform but also between each liganded state of each isoform could potentially aid our understanding of the role of each ligand and the role of this interface in each ligand dependent scenario.

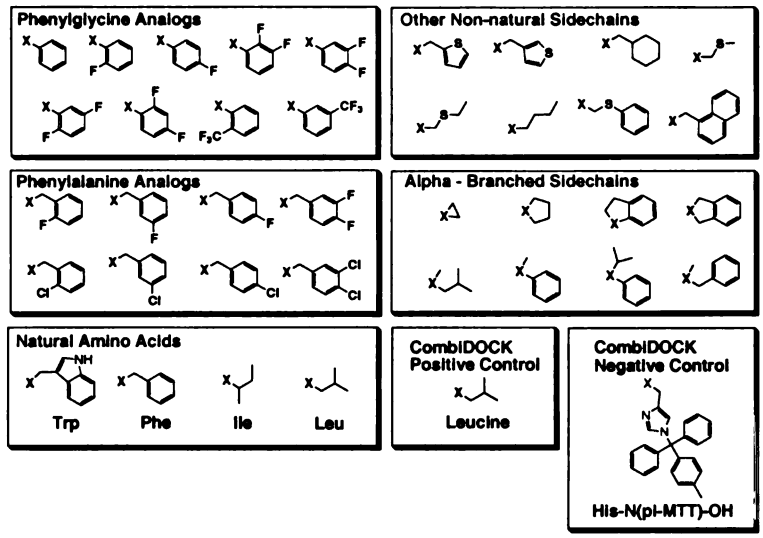
Intrigued by this idea, we wanted to explore the hypothesis that different agonist ligands, which are known to modulate the ER isoforms receptor to recruit SRC2-2 peptide, may create subtle differences within the LXXLL leucine binding pocket of ER isoforms that are beyond the reach of the natural leucine sidechains yet are significant enough to be identified by the non-natural chemistry we have introduced into our inhibitor library to reach farther into the binding pocket. To explore this idea we chose to examine subtle differences in the leucine binding pockets of hERα and hERβ in the presence of different ligands, including a synthetic estrogen analogue diethylstilbesterol

(DES), and the phytoestrogen genistein (Gen). Both ligands have been identified as partial agonists and have been shown to recruit SRC2-2 for both ER isoforms. Previouse reports of the effects on ERα and ERβ interactions with SRC peptides in the presence of these ligands demonstrates that while there were changes in the recruitment of the SRC NIDs [87] there was little change in the recruitment of the individual boxes.[76] ERα and ERβ have been determined to recruit SRC2 in the presence of E_2, DES and Gen with similar dissociation equilibria.[76] But while each of the ligands does not significantly perturb the binding of the the native LXXLL NR box motifs, it is possible that the hydrophobic pocket is modulated in such a way to maintain the binding of the three leucines but that other non-natural leucine mimetics would be significantly effected by the ligand dependent allosteric modulation of the pocket.

To examine this hypothesis we decided to evaluate the structural interface of the same receptor, hERα, in the presence of two different ligands DES and E_2 (DES•hERα•SRC2-2,[18] and E_2•hERα•SRC2-3[66]) to identify any subtle differences in the interfaces. Such a comparison of the same receptor with two different ligands and different SRC2 peptides, reveals that, despite being the same receptor, yet again there are significant differences within the pocket (Figure 3-1CD). Looking at the cross-section of each pocket we see topography of the receptor surface adjacent to both L_3 and L_1 of the SRC2 peptide. These differences were volumetrically determined to be a change of 401 $Å^3$ in both structures of hERα, indicating that there is not a change in the overall volume of the pocket but rather a change in the location of solvent accessible volume (Figure 2G). This last comparison is problematic due to the fact that there are multiple differences in the crystallization partners of hERα and we do not know if these differences are due to the change in the ligand or the SRC2 NR box peptide sequence. Yet given the large degree of differences in and around the SRC NR box recognition pocket of the nuclear receptors, it is possible that these interfaces are more dynamic than the static views that the crystal structures have provided. For example, the ability of the ligand chemistry to alter the surface may be more subtle than the gross structural effects on the position of helix 12 and be able to select for different non-natural side chains within this pocket.

Figure 3-2. Targeted diversity elements included in proteomimetic library as chosen by CombiDOCK.

This subset of amino acids was hydrophobic, exhibited good shape complimentarity to the leucine binding pockets, and were structurally diverse. Amino acid sidechains shown with the Cα designated as X. Diversity reagents were prepared as Fmoc protected α-amino acids and incorporated into the Proteomimetic library as individual substitutions at each of the leucine positions. Leucine and His(MTT) were the combiDOCK controls.



CombiDOCK Analysis

Computationally DOCKing to the receptor surfaces of hTRβ and hERα selected for not only leucine but other natural and non-natural sidechains (Figure 3-2) and there were differences in the sidecahins that were selected between the receptors. Final structures were sorted based on lowest energy conformations (Figure 3-3). Each top-scoring compound was evaluated visually to ensure that the molecules were minimized into a reasonable conformation in the receptor pocket (Sybyl).™ This final subset of molecules was then re-evaluated with a dynamic scaffold and were determined to have similar relative scores and final energy minimized orientations. Interestingly, some differences were seen between the fixed and dynamic scaffold in the selection of sidechains. For example, with T_3•hTRβ•SRC2-2 and DES•hERα•SRC2-2 the fixed scaffold tended to select residues such as phenylglycines and phenylglycine analogs (Figure. 3-3 D,E,F) that were shorter and/or could obtain orientations that allowed the overall helix to bind more closely to the receptor structure. On the other hand, the dynamic scaffold tended to select for residues such as phenylalanines and related analogues (Figure. 3-3 G,H,I) that forced the overall helix to bind in an orientation farther from the receptor surface. We included both subsets of side-chains in the targeted library.

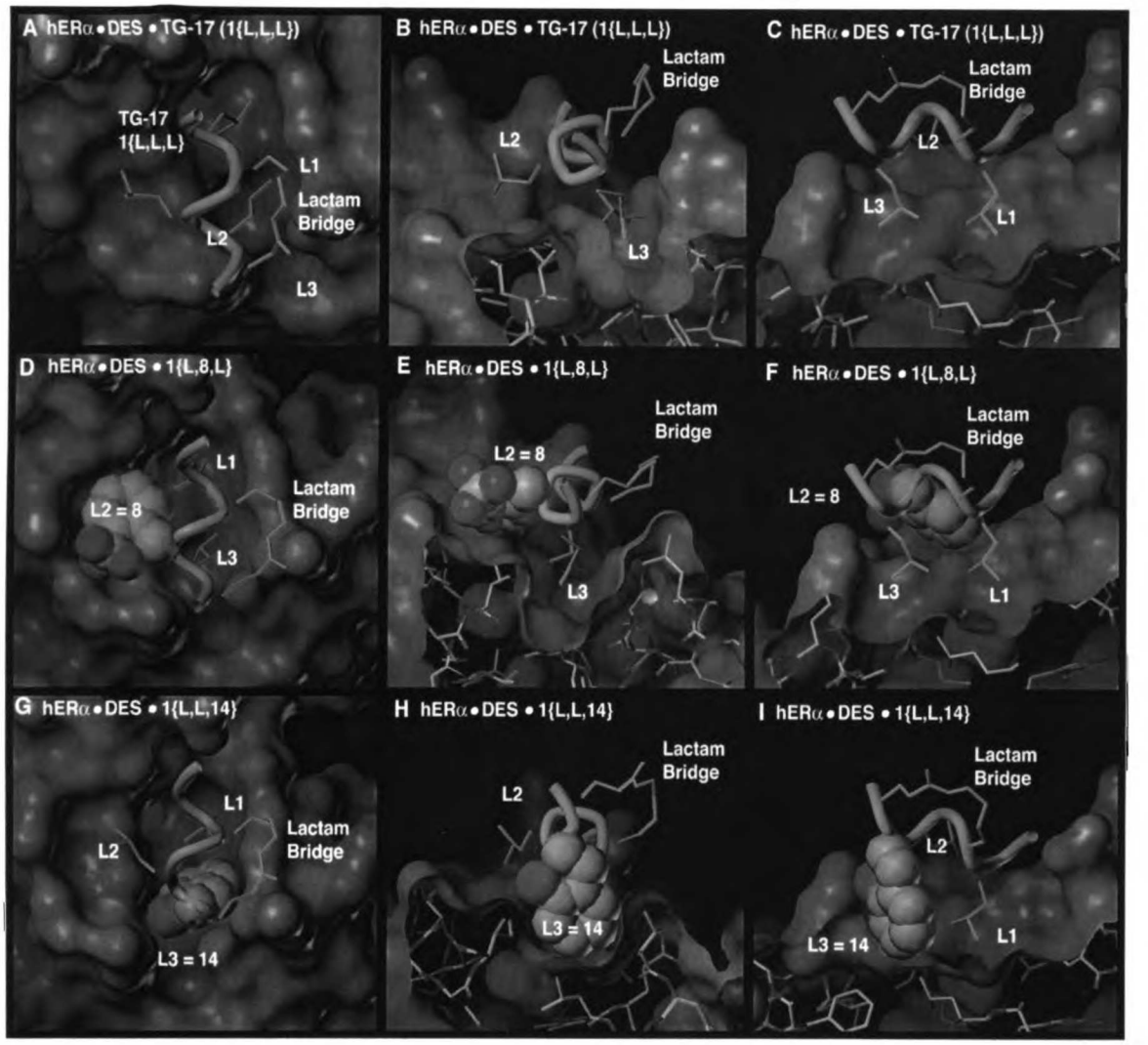


Figure 3-3. Representative CombiDOCK structures

Representative CombiDOCK structures of the proteomimetic scaffold and two inhibitors. Panels (**A-C**) show the energy minimized scaffold structure TG-17, **1{L,L,L}** composed of the thirteen amino acid SRC2-2 peptide sequence with a lactam bridge constraint from the i to i+4 position DOCKed to hERα•DES from the top (**A**) and two cross-sections (**B-C**). *In silico* combinatorial evaluation of libraries sidechains bound to the **1{X,X,X}** scaffold selected for compounds that were likely to bind to either hTRβ or hERα. Panel (**D-F**) show the energy minimized structure of the hERα•DES inhibitor **1{L,8,L}** with Ortho-Trifluoromethyl-phenylglycine in the second leucine position bound to hERα•DES. **1{L,8,L}** scored the highest in the docking studies with both receptors hERα•DES and hTRβ•T_3. Panels (**G-I**) show the another inhibitor **1{L,L,14}** with 2-Ortho-Fluoro-Phenylalanine in the L_3 position DOCKed to hERα•DES and which scored better with hERα•DES than with hTRβ•T_3. Each of the DOCKed inhibitor structures were similar in orientation to that of the crystallographic coordinates of SRC2-2 Docked to either of the receptors. The RMSD values were: **1{L,L,L}** at <0.1Å; **1{L,8,L}** at <0.2Å; **1{L,L,14}** at <0.6Å. These structures were combinatorially DOCKed based on size shape complimentarity, with a constrained helical scaffold that was allowed to deviate from the original minimized crystal coordinates by 2.5Å utilizing CombiDOCK[88]. One position was substituted at a time while holding the remaining two positions as Leucines.

Conclusion

Computational studies revealed subtle differences between two NR box recognition interfaces, ER and TR. Additionally, differences were seen within the same receptor isoform ERα, depending on what ligand (E_2 or DES) and/or NR box peptide (SRC2-2 or SRC2-3) was bound. Based on some of these differences we were able to select for a wide variety of non-natural residues at each of the leucine positions including several natural sidechains that nature had not selected for, including phenylalanine and tryptophan. 37 residues (Figure 3-2) were selected for incorporation at each leucine position within the lead scaffold.

While this approach is limited by a lack of consideration for dynamics in receptor surface structure, electrostatics, and solvation, it was useful in reducing the size of the target library to members that were likely to bind to the receptor in a conformation that is similar to that seen in the crystal structures. These other parameters can be considered, however, with tremendous increase in processor time. Weather or not these residues could be successfully incorporated into the scaffold molecule was still not clear. Additionally, we could not clearly predict if these differences within the sidechains would be significant enough to select between the receptors.

Methods

Computational Design - CombiDOCK

Design of the library consists of selecting the correct amino acids for inclusion within the synthesized library. Some 3000+ non-natural amino acids from the publicly available chemicals directory (ACD)[89] were identified (MDL-ISIS/Draw 2.4) are screened *in silico* by DOCKing to the coactivator binding surfaces of the relevant receptor structures. For example, the SRC2-2 peptide from each of the co-crystal structures, T_3•hTRβ•SRC2-2 and DES•hERα•SRC2-2, was utilized as the scaffold for the presentation of non-natural amino acid leucine replacements to DOCK into a crystallographically defined negative image of each receptor. The SRC2-2 peptide of each of the protein complex crystal structures was edited to remove the distal atoms of the three leucine side chains to Cβ, leaving the rest of the peptide intact. Each of the conformers generated from potential

amino acids was then attached to the SRC2-2 structure by anchoring the components according to the Cα and Cβ carbon coordinates for each leucine. The compounds were then evaluated for binding using CombiDOCK in two ways: 1) with each of the leucine positions varied combinatorially to generate 5896^3 possible combinations and, 2) with each position varied separately while holding the other two positions as leucine to generate 5896 x 3 possible compounds. The SRC2-2 peptide backbone was held structurally rigid to maintain the alpha-helical structure while each compound was energy minimized in the receptor pocket to maximize the size-shape complimentarity.

The distance dependence van der Waals docking parameter followed the following equation where E_{VDW} is the intermolecular interaction energy; C and D are calculated coefficients; e is the well depth of interaction energy; R is the van der Waals radius of the atoms; r is the distance between atoms; a and b are the van der Waals repulsion and attraction components.

$$E_{VDW} = Ce\ (2R/r)^a - De(2R/r)^b$$

PDB structures T_3•hTRβ•SRC2 (PDB: 1BSX) and DES•hERα•SRC2 (PDB: 3ERD) were utilized for computational evaluation. The ACD was searched for α-amino acid molecules (AA) with point of diversity at the Cα position. (ISIS/Base)™. AA structures were converted to SMILES strings with unique FCD numbering[90] (*Merlin*) A Daylight fingerprint [91]with fixed length and low density was generated for each molecule; selections were based on: hydrophobic side-chains, ClogP values; size (MW < 500) and reactive functional groups. AA were hierarchically clustered (*Cluster*) to remove redundant side-chains and groups that fill the same conformational space based on the radial overlap between spheres of each molecule. A nxn distance matrix was calculated for the similarity of the compounds based on nearest neighbors command. The similarity cutoff was changed (sim.cutoff = 0.78) until a final set of < 500 molecules were defined to be different and chosen for the library. Each amino acid was then reduced to its side chain, including the alpha carbon (Cα) (*Sybyl*)®.[92] Anchoring numbers were defined as Cα = 3, Cβ = 2, Cγ = 1 and applied to each molecule within the PDB file with an in house script.[93] All energetically refined conformational rotamers were generated

(*Omega*).[Openeye Scientific Software #138] Input global parameters were: RMS cut-off 0.8; energy window 6.0; maximum conformational output number was evaluated at different levels and a final range of 10-100 per molecule was utilized; maximum population size 200,000; maximum rotors per molecule 17. This generated a torlib.txt torsion library. The heavy atoms from the SRC2-2 peptide were then added to the side chain library assembly with anchoring set as: N = 3; Cα = 2; Cβ = 1; for each of the leucine positions (*Sybyl)*®. A final set of molecules was output to a fragment molecule file (fragment.mol2) to be utilized by the CombiDOCK program discussed below.

Each receptor coactivator binding pocket was prepared creating a negative image of the surface by Sphgen under default setting parameters. GRID generated a contact grid with default parameters set and the following parameters: (grid spacing 0.3Å; contact cut-off 4.5Å; energy cut-off distance 10Å; bump overlap 0.75Å). No electrostatics were applied. The OMEGA input global parameters were: RMS cut-off 0.8; energy window 6.0; maximum conformational output number 10 per molecule; maximum population size 200,000; maximum rotors per molecule 17; a torlib.txt torsion library. The OMEGA active site mode parameters were: energy window 10.0; chemscore grid energy grid resolution of 0.3; maximum random translation 1.0; maximum random rotation 30.0; number of anchor positions 200; output number of conformers 10. CombiDOCK scoring parameters: no old grids; bump filter set with no maximum; contact score set; no chemical score; energy scoring and interpolation set; a Van der Waals scale range of 1.0 to 2.0Å during different evaluations; no electrostatic scale; no energy maximum. The combiDOCK minimization parameters were as follows: energy minimization set; probe minimization set; the orientation and position of the peptide was allowed to deviate from the original starting position by a maximum RMSD of 1.5 angstroms; an energy convergence of 1.0; no torsion minimization; a maximum of 100 iterations; initial translation tolerance of 1.0; initial rotation tolerance of 5.0. CombiDOCK matching parameters were set to: matching of ligand centers; 3 nodes minimum; 6 nodes maximum; distance tolerance of 0.2Å to 0.8Å were set for different evaluations; distance minimum was 1.0 to 2.0; 100000 maximum matches; no match ratio minimum, check degeneracy, critical spheres nor chemical matches were performed. CombiDOCK combinatorial parameters: three total sites (corresponding to L_1, L_2 and L_3, or each

leucine); check clashes set; clash dependents were set for each position relative to each other (1 vs. 2, 1 vs. 3, 2 vs. 3); all 5896 ligands were placed at each position in a combinatorial or sequential manner (see below); greedy conformations were varied from 10-100; fragments were merged; 20 maximum anchor torsions; pre-computed clashes were set to be evaluated; greedy scaffolds were varied from 10-500; and one probe was set at each position for initial scaffold configuration; the scaffold was minimized with methyl groups at each position. Side chain spheres that violated van der Waals distances with the SRC2-2 peptide or each other were automatically discarded due to clash violations before docking proceeded. The scaffold was minimized to the receptor site with methyl groups and then each sidechain compound was added and evaluated. General parameters: ligand was scored and oriented in a combinatorial or non-combinatorial fashion; no rule of five; multiple orientations evaluated. Active site mode parameters were: energy window 10.0; chemscore grid energy grid resolution of 0.3; maximum random translation 1.0; maximum random rotation 30.0; number of anchor positions 200; output number of conformers 10. The Van der Waals component uses a modification of the Leonard-Jones 6-12 attraction and repulsion components with distance dependence. Contact evaluates and penalizes (user defined) for making contacts within 4.5Å between two non-hydrogen atoms. CombiDOCK reads in the receptor files (receptor.mol, .grd, .sph) and the fragment (.mol) library file. A Van der Waals definition file (vdw.def) is also read in conjunction with grid. Chemical Scoring is an option that we did not utilize but incorporates empirical theories and constraints, and a modifications of the VDW terms. The combiDOCK general parameters were as follows: ligand was scored and oriented in a combinatorial or non combinatorial fashion; no rule of five; with multiple orientations.

Final structures were sorted based on lowest energy conformations. Each top-scoring compound was evaluated visually to ensure that the molecules were minimized into a reasonable conformation in the receptor pocket (Sybyl).™ (Depending on the parameter settings and scoring constraints, molecules could dock into a conformation that was dramatically different from that of the crystal structure. For example, a large polycyclic mono-methyl-trytyl protecting group on the Histidine (MTT) (Figure 3-2) was used as a negative control which, when given the appropriate parameter settings, would score well,

however, the entire scaffold would be shifted outside the pocket). A positive control was the natural, Leucine sidechain. This final subset of molecules was then reevaluated with a dynamic scaffold (CombiDOCK general parameters, orient ligand – yes) and allowed to deviate in orientation by three angstroms from the starting position by changing the CombiDOCK minimization parameters (max_rmsd, initial translation and rotation) and were determined to have similar relative scores and final energy minimized orientations. Often, different results will be obtained depending upon whether or not the scaffold is fixed during the CombiDOCK run. For example, with T_3•hTRβ•SRC2-2 and DES•hERα•SRC2-2, the fixed scaffold tended to select residues such as phenylglycines and phenylglycine analogs (Figure 3-2) that were shorter and/or could obtain orientations that allowed the overall helix to bind more closely to the receptor structure. On the other hand, the dynamic scaffold tended to select for residues such as phenylalanines and related analogues that forced the overall helix to bind in an orientation farther from the receptor surface. It is beneficial to include both subsets of side-chains in the targeted library.

Volumetric Analysis

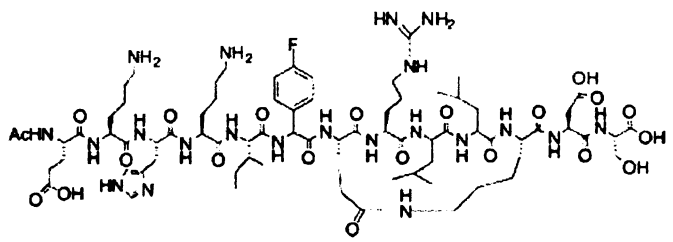
Publicly available structures, T_3•hTRβ•SRC2-2 (PDB: 1BSX), E_2•hERα•SRC2-3 (PDB: 1GWR) and DES•hERα•SRC2-2 (PDB: 3ERD) were utilized for computational evaluation. A perl script was written to identify solvent accessible volume on the surfaces of protein.[94] The PDB structures were aligned by the Cα and Cβ positions of the three leucines (or two leucines and a tyrosine for 1GWR) with RMSD <0.01A of the SRC2 coactivator peptide. The coactivator peptide was removed from the structure and a box of 4x4x4 angstroms^3 in dimension was defined with a center defined in PDB coordinate space as the center of the binding pocket. Spheres filled the pockets of both receptors. Those that came in contact with the receptor surface of a given receptor were removed. The sphere sets were compared between receptors and those that were identical in both receptors were deleted, while those that were different between receptors were defined as a receptor specific difference indicating a 2.2 angstrom radius sphere fit in one structure and not another.

Acknowledgements

We would like to thank Dr. Bea Darimont for critical assistance early in the development of this project and Dr. Yoko Shibata for the expression of hERα. TRG was funded by the Department of Defense Breast Cancer Research Fund #DAMD17-00-1-0191. We thank the Sidney Kimmel Foundation for Cancer Research, the HHMI Research Resources Program grant #76296-549901, the Academic Senate of University of California at San Francisco, NIH #DK58080, and the Sandler Foundation for financial support. We would like to dedicate this chapter to the memory of a great mentor and colleague, Dr. Peter Kollman. His thoughts and work will always be with us; and he is dearly missed

Chapter IV Novel Selective Inhibitors of the Interaction of Individual Nuclear Hormone Receptors with a Mutually Shared Steroid Receptor Coactivator 2

Timothy R. Geistlinger; Guy, R. K., Novel Selective Inhibitors of the Interaction of Individual Nuclear Hormone Receptors and a Mutually Shared - Steroid Receptor Coactivator 2, *J. Am. Chem. Soc.*; **2003;** *125*(23) pp 6852 – 6853



Timothy R. Geistlinger.; R. Kiplin Guy, Steroid Receptor Coactivator Proteomimetics, *Methods In Enzymology*, **2003**, Vol. 364, Chapter 13, pp 223-246

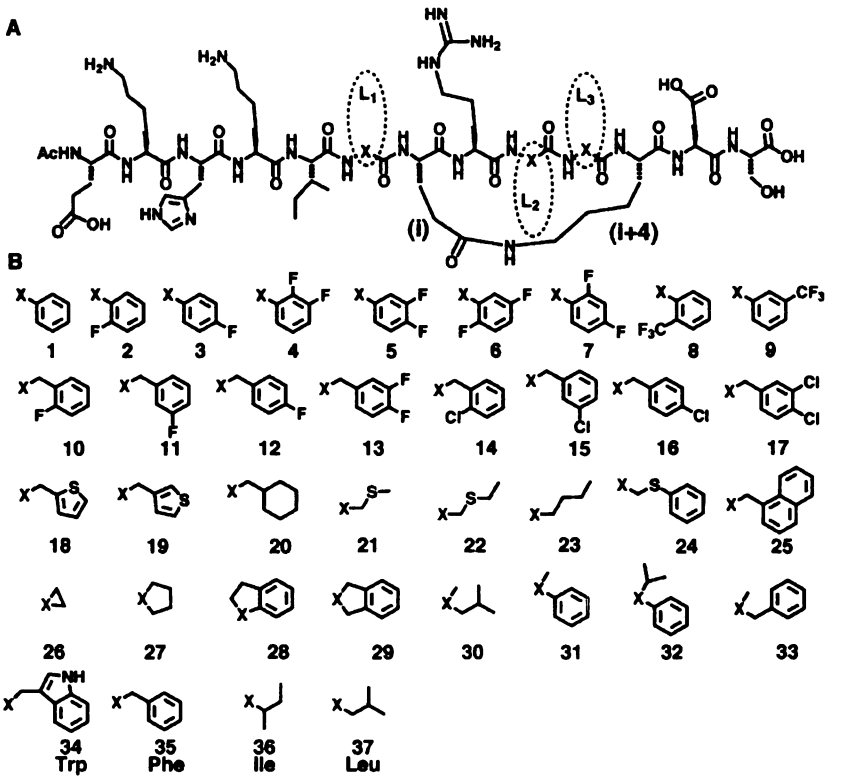
Introduction

Based on our computational evaluation we determined that while hERα, and hTRβ interact with the SRC2-2 peptide through similar surfaces [7,18,67,72,73,75] if we apply the non-natural sidechains to the identified scaffold, Chapter II, it may be possible to discover novel inhibitors of coactivator binding that are selective to particular NR. We designed a library of α-helical proteomimetics, constrained by a macrolactam at positions E^{691} and K^{695}, that mimic SRC2-2 {c(E^{691}-K^{695})Ac-^{685}EKHKI**L_1**ER**$L_2$$L_3$**KDS^{697}-COOH} (**1{37,37,37}**) with non-natural amino acids replacing the leucines at positions **L_1, L_2,** and **L_3**. The selection of particular non-natural side chains for this library was governed by computational DOCKing (CombiDOCK)[95] methods (see Chapter II) using hERα•E2•SRC2-2, and hTRβ•SRC2-2 X-ray crystal structures. These non-natural residues were synthesized and chemically substituted at each leucine position (Figure 4-1). The resulting molecules were screened *in vitro* against the coactivator binding surfaces of the receptors.

Results

Computationally, the NR selected for similar residues at each position except that hERα tolerated larger and Cα branched residues (Figure. 4-1B # 24, 25, 27-34). Studies were carried out in which each leucine was replaced individually without changing any other amino acids and in which all three leucines were simultaneously varied. As the CombiDOCK studies did not reveal significant cooperativity between the diversity positions, the targeted library was limited to a single substitution in each member.

Figure 4-1. Scaffold and Library of Diversity Elements



Panel A) The structure of the proteomimetic library scaffold **1** with a lactam constraint at positions i to i+4, and diversity positions indicated by L_1, L_2, and L_3. 1{37,37,37}: L1 = L2 = L3 = LEU. Panel B) The targeted diversity elements included in proteomimetic library as chosen by CombiDOCK. Library was produced by substitution of a single L_x for each member with other positions being left as leucine. Amino acid side chains shown with the Cα designated as X. Superimposing the X on the scaffold with the X on the side chain reassembles individual library members.

The designed library was synthesized in parallel on solid support utilizing Fmoc chemistry as previously described for constrained SRC2-2 mimetic **1{37,37,37}**.[81] In general, the synthesis proceeded smoothly to yield milligram quantities of each compound after RP-HPLC purification (Table 4-1). Many of the Cα branched compounds proved to be more difficult to synthesize and purify, particularly during macrolactam formation. The final proteomimetic library contained 87 of 111 targeted compounds composed of 25, 30, and 32 individual substitutions, respectively, at positions L_1, L_2, and L_3 (Figure. 4-1A). The purity and identity of all compounds were confirmed by LCMS and HR-MALDI-TOF MS, respectively (Figure 4-5,6,7).

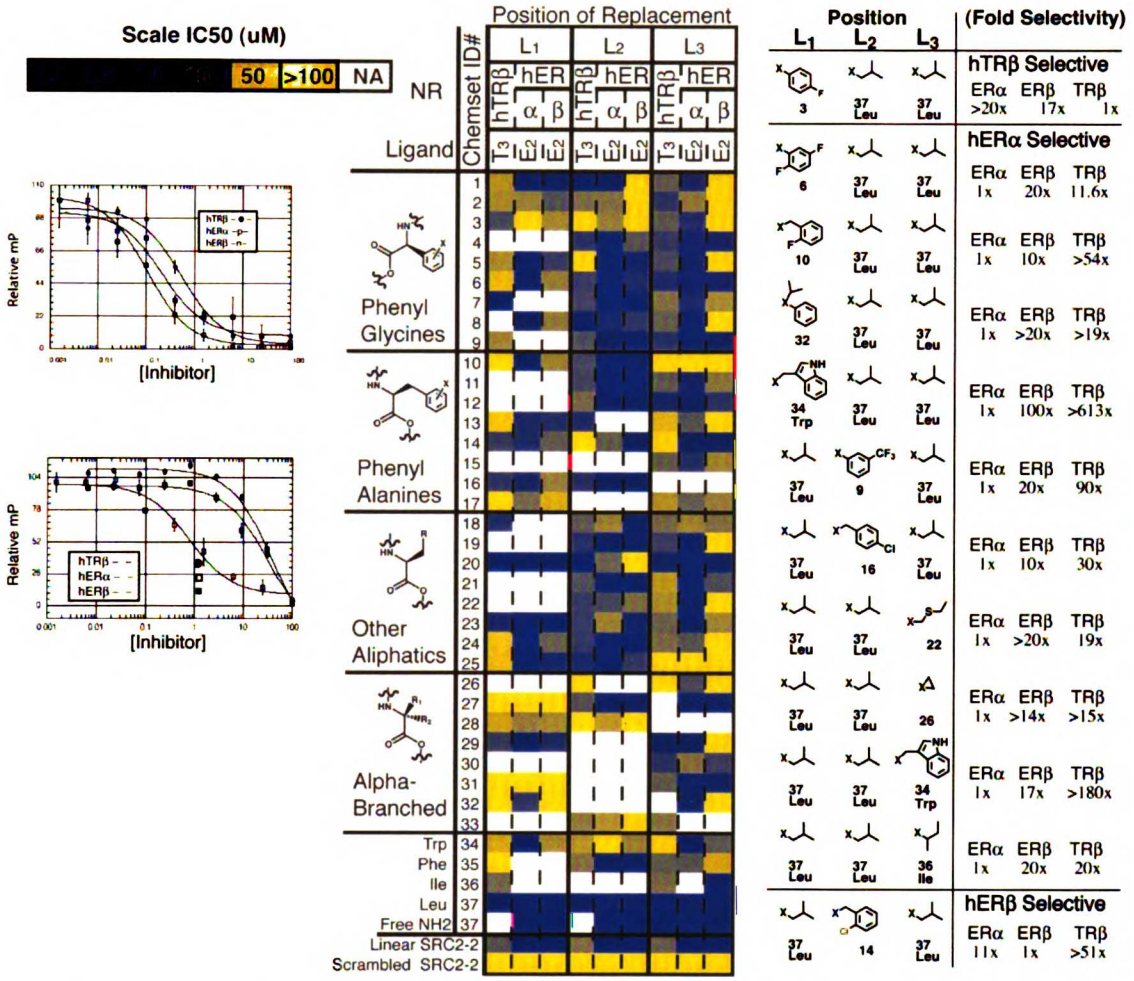


Figure 4-2. Competition Assay Results of a library of proteomimetics of SRC2-2..

Inhibitory activity and selectivity profile of the SRC2-2 proteomimetics. The equilibrium 50% inhibitory concentration (IC_{50}) for each library member for competition of the SRC2-2 peptide from each NR (hTRβ, hERα, or hERβ) as determined by *in vitro* fluorescence polarization assays. Activity is represented as a colorimetric scale, with light blue indicating an IC_{50} >100μM, dark blue indicating an IC_{50} of <100nM, and gradations of color between the two indicating intermediate IC_{50} values. White boxes denote compounds whose synthesis was not achieved. Individual non-natural amino acids are arrayed on the Y-axis, and numbered according to Figure. 4-1B. The X-axis depicts the position of non-natural amino acid substitution (L_1, L_2, or L_3) and the NR tested. The structures of twelve library members that are specific to one NR due to individual side chain substitutions at one position. Selectivity among the tested NR is indicated on the right side with the fold decrease in the IC_{50} relative to the IC_{50} against the receptor for which the compound is most selective.

The ability of the proteomimetics to compete with SRC2-2 peptide for binding to the NR was assessed using fluorescence polarization (FP) equilibrium competition assays.[81] These studies revealed that the library members exhibited a large range of inhibitory ability for blocking the binding of the SRC2-2 peptide to the NR (Fig. 4-2). As hypothesized from the DOCKing studies the majority of phenylglycine and phenylalanine analogs allowed for effective competition with 71 of 87 compounds giving IC_{50}'s that

were equivalent to or better than the SRC2-2 peptide with one or more NR. The high level of success in picking compounds with competitive ability validates the use of the CombiDOCK methodology for the design of inhibitors of protein-protein interactions in addition to its conventional use in design of enzyme inhibitors.

Discussion

The testing of the inhibitors identified the first selective proteomimetics (Figure 4-2) that take advantage of differences between the leucine binding pockets of E_2•hERα, E_2•hERβ and T_3•hTRβ. A total of 12 compounds were at least 10 fold selective for binding to hERα in preference to hTRβ or hERβ. Of these, a number were more than 20 fold selective and one, **1{34,37,37}**, more than 600 fold selective for hERα. Surprisingly, only one compound, **1{3,37,37}**, was selective for hTRβ and one, **1{37,14,37}**, for hERβ. Strikingly, two natural amino acids, tryptophan and isoleucine provide high levels of selectivity for hERα. Previous studies using genetic selection from random peptide libraries failed to reveal this trend instead selecting exclusively for leucines within the $L_1XXL_2L_3$ motif while evolving differences in flanking sequence.[8,77]

The SRC binding pockets of NR have evolved to bind to a simple hydrophobic $L_1XXL_2L_3$ consensus motif while relying upon differences in SRC structure flanking the NR box to convey selectivity *in vivo*. Our studies reveal that the NR box binding pockets of the NR contain significant differences in shape and electrostatics that allow competitive inhibitors that mimic the NR box to act selectively on one NR. The fact that discrimination between NR can be achieved solely by manipulation of the side chains inside the small $L_1XXL_2L_3$ motif implies that a suitable small molecule could achieve the same result. Thus, targeting this site may prove to be a general method for producing selective modulators of NR function.

These results imply that the inhibitor selectivities are derived from within the three leucine recognition pocket and indicate that there may be more subtle differences within the pocket which is permitting the NR specific selection of these inhibitors. They also may suggest that our conventional, static view of this interaction may be limited and requires further investigation into the mechanisms of action behind this interaction. Our computational analysis of different NR hydrophobic SRC recognition pockets, as

described in chapter three, demonstrates that there are significant differences in amino acid composition of the residues making up the hydrophobic pockets on each receptor which create differing electrostatic potentials and topologies throughout the pocket. And when looking at the differences in hERα•E_2 and hERα•DES we have identified noticeable differences. Ligands, Estrogen, and estrogen-like ligands, for example, function by binding with high affinity to the ligand binding domain (LBD) defining the hydrophobic core of the domain as described in chapter I. The agonist ligands allosterically define the overall structure creating a conformational change in the LBD to promote coactivators recruitment. Because these effects are significant, we wondered if it would be possible to identify sub-sites within the three leucine coactivator binding pocket of a given receptor which change in a ligand dependent manner yet have no apparent effect on the three leucines, ie... agonist ligands that still recruit the LXXLL peptides, but modulates the pocket to select for different non-natural sidechains of the inhibitors. Until now the focus of ligand allosteric effects has been on gross structural changes like the relative modulation of the helix 12 position within the LBD and its ability to create or conceal the SRC binding pocket to act as an agonist or an antagonist, respectively. Each ligand seems to have a different effect on the relative equilibrium between these states and between estrogen receptor isoforms. When viewed crystallographically or evaluated biochemically, however, the differences are subtle and unclear as to how these differences would generate significantly different biological effects. Generating inhibitors to the protein-protein interaction between the SRC and each ER isoform that were selective not only to each isoform but also between each liganded state of each isoform could potentially aid our understanding of the role of each ligand and the role of this interface in each ligand dependent scenario.

Materials and Methods

All computational evaluation was performed on a Silicon Graphics Inc., Octane, IRIX 6.5, operating system. CombiDOCK(Kuntz Laboratory, UCSF). Wang resin and natural, alpha amino acids were purchased from [s] NovaBiochem, San Diego, CA, and non-natural alpha amino acids were from various sources (see **Table 1**). ABCR GmbH & Co. KG,[t] Karlsruhe, Germany, Peptech Corp,[u] Cambridge, MA, Advanced Chemtech,[v]

Louisville, KY, Chem Impex,[w] Wood Dale, IL, Bachem [x] California, Torrance, CA, ACROS Organics,[y] Merris Plains, NJ, Salor,[z] Milwaukee, WI. Oregon Green 488 succinimidyl ester (MP# 0-6147) was purchased through Molecular Probes, Eugene OR. Robbins Block (Robbins Scientific Co., Sunnyvale, CA). TetrakisTriphenyl-phosphine-Palladium (0) [34] (STREM, Newburyport, MA), [34], HBTU (Advanced Chemtech, Luoisville, KY), PyBOP (NovaBiochem, San Diego, CA), HATU (Perseptive Biosystems, Warrington, England), Genevac HT-4 Series II Evaporating System (Genevac Limited, Ipswich, UK). Parallex Flex, BIOTAGE HPLC (CombichemLabs, Patterson CA); RP-C_{18}, Reliasil, BDX-C18, 5μm, 50x21mm Preperative HPLC Column, (Column Engineering, Ontario, CA); DE-STR-MALDI (Precision Biosystems, Framingham, MA). MALDI-Quality Alpha-cyano-4-Hydroxycinnamic Acid in Methanol, (Hewlett Packard). *1D ^1H NMR (400MHz)* was performed on a Varian Inova, (Varian, CA); Circular Dichroism Studies were performed on a Jasco J-715 Spectrometer (Japan); Protein expression required the following: Thyroid hormone [(3,3',5-triiodo-L-thyronine (T3; Sigma)], Estrogen (Sigma), Protein chromatography - HPLC (TSK-phenyl column (TosoHaas, Philadelphia, PA)) and NAP gel filtration (Pharmacia), BCA assay reagent (Pierce). MALDI-TOF Mass spectrometry (Voyager-DE, PerSeptive Biosystems). All fluorescence spectral data were acquired on an LJL Biosystems Analyst AD system, Criterion™ Software (LJL Biosystems, Inc., Sunnyvale, CA).

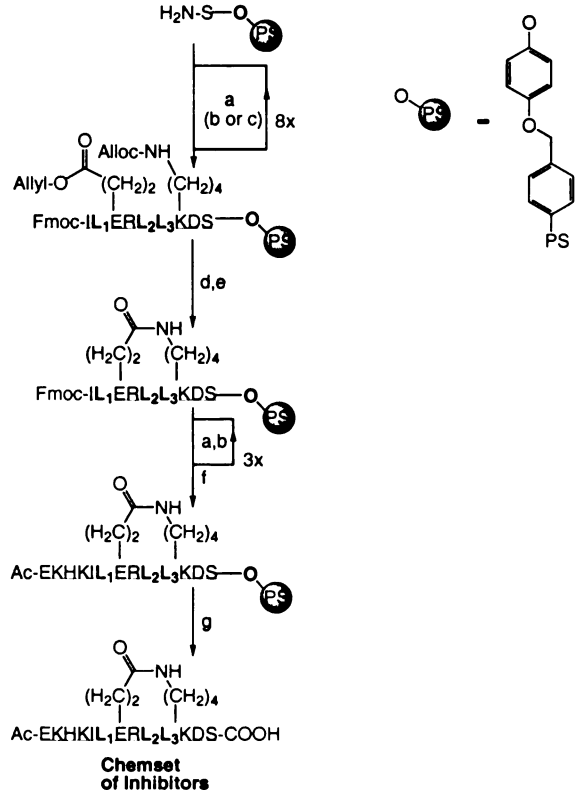


Figure 4-3. Chemset synthesis

Underline indicates a single letter amino acid code. Bold letters (L_1, L_2, L_3) indicate a position of diversity in place of the conserved leucines a) 20% piperidine in DMF, 10 min, rt, repeated twice per cycle, b) {natural amino acids} Fmoc amino acid (3 eq.) HBTU (3 eq), DIEA (3-5 eq) in DMF, 0.5-2 hrs, rt, repeated twice per cycle; c) {non-natural amino acids} HBTU (1 eq), PyBOP (1 eq) , HATU (1 eq), TFFH (1 eq), DIEA (5 eq) in DMF, 2 hrs, rt, repeated three times; d) Pd(Ph3P)4 (0.02 eq), N,N-dimethyl barbituric acid (3 eq), DCM, rt, 2 hrs, repeated three times; e) PyBOP (1 eq) TFFH (1 eq), HOBt (1 eq), DIEA (3-5 eq); in 1:1 NMP:DMF, shaken, 2 hrs, rt, repeated twice. f) Ac_2O (3 eq), DIEA (3 eq) 30 min, rt; g) 1% H20, 2%

Phenol, 2% Thioanisol, in TFA, rt, 3 hrs.

Library Construction

Orthogonally protected peptides for the chemset were synthesized on solid support in Robbins Blocks™ (48 wells / block) with 100mg resin/well with a resin loading of 0.48mmol/g. Each synthesis involved the following series of steps: A) Addition of each additional amino acid followed a cycle of: 1) removal of the amino terminal Fmoc protecting group (20% piperidine in DMF, 10 min, rt, repeated twice per cycle); 2) Addition of the next Fmoc amino acid (3 eq.) (HBTU (3 eq), DIEA (3-5 eq) in DMF, 0.5-2 hrs, rt, repeated twice per cycle); B) non-natural amino acids were activated using a combination of coupling agents (HBTU (1 eq), PyBOP (1 eq) , HATU (1 eq), TFFH (1 eq), DIEA (5 eq) in DMF, 2 hrs, rt, repeated three times. The protected peptide was assembled to the isoleucine[689] residue (Fmoc-NH-[689]IL_1E(allyl)RL_2L_3K(alloc)DS-Resin) two residues amino terminal to the orthogonally protected acidic residue E(allyl)[691]. It became apparent that performing this step earlier improved yields by increasing intra-peptide lactam formation and reduced inter-peptide lactam formation on the resin. Formation of the lactam bridging group followed the sequence of: 1) Orthogonal deprotection of allyl and alloc protecting groups via $(Pd(Ph_3P)_4$ (0.02 eq), N,N-dimethyl barbituric acid (3 eq), DCM, rt, 2 hrs, repeated twice. Cyclization (PyBOP (1 eq) / TFFH (1 eq), HOBt (1 eq), DIEA (3-5 eq); in 1:1 NMP:DMF, shaken, 2 hrs, rt, repeated twice). Both were monitored via the Kaiser test at each step to evaluate progress. After each deprotection step a positive blue color was noted. This color was reduced after each round of lactam coupling. The cycle was repeated three times for each of the constrained peptides. When there was no change in the Kaiser test after each round of deprotection the Fmoc peptide synthesis was continued until the end of the peptide and the resulting amino terminus was acetylated (Ac_2O (3 eq), DIEA (3 eq) 30 min, rt). Final steps followed the sequence of: 1) cleavage from the resin and concomitant side chain deprotection (1% H_20, 2% Phenol, 2% Thioanisol, in TFA, rt, 3 hrs.); 2) ether precipitation; 3) HPLC purification. (Figure 4-3)[81]. Compound identity was confirmed by UV spectroscopy, MALDI-TOF MS, and ESI Quadrapole MS. Several of the compounds required multiple rounds of RP-HPLC with different conditions to achieve suitable purity. Final library purity was assessed using RP-HPLC-ESI-TOF with a different

column, mobile phase, and temperature, verifying that most library members had a purity greater than 80% with the majority being >90% as determined by the lowest purity value of either the 215nm UV absorption spectrum or TIC (Figure 4-4,5,6). Compounds that were below these values are indicated in Table 4-1 with the major contaminant being the intermolecular cross-linked peptide. An example of a compound with low purity is **26,37,37** shown (Figure 4-5). HR-ESI-TOF and HR-MALDI-TOF experiments confirmed that all the compounds met the expected mass.

	Yield (%Thoretical)			Calc Mass	Measured Mass (m/z) MALDI-TOF			Purity %		
Compounds	L_1	L_2	L_3	m/z	L_1	L_2	L_3	L_1	L_2	L_3
1[v]	1.51	1.82	3.33	1652.89	[a]1652.90	[a]1652.90	[a]1652.90	95	95	95
2[t]	0.60	0.60	4.49	1670.88	[a]1670.91	[a]1670.94	1671.07	95	90	95
3[t]	0.30	3.44	2.24	1670.89	[a]1670.57	1670.66	1670.69	95	90	95
4[t]	0.00	1.63	3.26	1688.87	-	[a]1688.82	1689.02	-	80	95
5[t]	1.63	0.30	5.92	1688.87	1689.02	[a]1688.99	1689.33	90	75	95
6[t]	1.92	1.63	6.51	1688.87	1688.99	1688.96	1688.9	90	80	95
7[t]	1.48	2.81	6.81	1688.87	1689.02	1689.02	1689.05	95	95	95
8[t]	0.00	5.96	3.63	1719.89	-	1720.8	1720.90	-	95	70
9[t]	0.58	10.32	1.31	1719.89	[a]1720.87	[a]1720.87	[a]1720.8	85	95	95
10[u]	1.19	12.61	4.30	1684.91	[a]1684.78	1684.88	1684.97	90	95	80
11[u]	0.00	5.49	6.38	1684.91	-	1685.13	1685.10	-	90	80
12[u]	0.00	2.67	4.45	1684.91	-	1684.82	1685.00	-	85	95
13[v]	2.20	3.67	1.17	1702.9	1702.91	1702.90	[a]1703.10	80	95	85
14[u]	3.67	0.88	15.43	1701.36	1701.04	1701.11	1701.08	90	95	95
15[u]	0.00	0.00	1.18	1701.36	-	-	[a]1701.19	-	-	90
16[u]	3.67	1.47	0.00	1701.36	1701.31	1701.35	-	90	95	-
17[u]	1.44	0.00	0.72	1735.80	[ac]	-	[ac]	85	-	95
18[w]	0.30	19.28	8.97	1672.95	[ac]	1673.81	1672.87	85	95	85
19[u]	0.30	2.39	5.08	1672.95	[‡]1672.41	1672.47	1672.51	80	95	95
20[s]	4.63	9.71	1.49	1672.97	1672.69	1672.89	1672.95	80	95	85
21[x]	0.00	8.09	3.67	1636.92		1636.87	1636.79		95	85
22[w]	0.00	1.51	5.91	1650.94		1650.86	1650.77		80	85
23[w]	2.91	4.59	1.68	1632.90	1633.02	1632.73	[c]	80	85	80
24[x]	2.80	0.59	6.47	1698.98	1699.09	1698.93	1698.3	95	95	95
25[u]	2.77	4.37	5.53	1716.98	1716.98	1716.9	1717.1	95	95	95
26[s]	37.43	9.36	5.77	1602.83	1603.05	1602.96	1603.02	[b]70	[b]65	[b]75
27[w]	1.69	1.07	0.00	1630.89	1630.91	1630.87		[b]60	85	
28[y]	7.30	1.94	0.00	1678.93	1678.29	1678.45		[b]70	[b]80	[b]70
29[y]	1.94	0.00	1.64	1678.93	1678.94		1678.9	[b]65		[b]70
30[x]	0.00	0.00	6.68	1646.93			[c]		[b]65	80
31[z]	0.60	0.00	0.30	1666.92	[a]1667.12		[a]1667.31	[b]70		[b]60
32[w]	6.64	0.00	0.44	1694.97	[c]		[ac]	[b]65		[b]60
33[w]	0.00	7.44	1.49	1680.95		1681.10	[a]1681.00		[b]80	[b]65

34[a]	3.37	10.55	5.72	1705.95	1705.76	1706.00	1705.89	95	95	95
35[a]	0.75	1.65	9.60	1666.92	[a]1667.28	1666.66	1666.7	90	95	85
36[a]	3.22	0.00	5.82	1632.90	1632.93		1632.9	85		95
37[a]	8.57	10.26	3.98	1632.90	1632.93	1632.93	1632.93	95	90	80
Constrained SRC2-2										
37[a]	3.03	40.91	21.21	1649.95	1649.93	1649.99	1649.96	95	95	95
Linear SRC2-2										
Scrambled SRC2-2	0.00	30.30	0.00	1649.95		1649.99			95	

Table 4-1. Yields and Characterization

Compounds are shown as chemset # and diversity element number. Substituted positions L_1, L_2 and L_3 are shown in respective columns L_1, L_2, L_3 and represent the individual compound data. [a] Compounds were analyzed in binding buffer by both MALDI-TOF and LC-ESI-MS. In cases of buffer saturation of signal, buffer was separated from the target compound by RP-HPLC and identifide by ESI-MS. [b]Major contaminant was the peptide dimer. [c]Samples were analyzed only by ESI-MS due to limiting amounts of material. [e]Compound appeared to break down over a two month time period during freeze thaw cycles in the buffer conditions. All compounds were stored at –80c. [s]

Preparation of Fmoc Protected Amino Acids Fmoc-X(1-9, 18,19,24,26-28,30-33)-OH by Fmoc-succinimidal amino protection

The non-natural amino acids were Fmoc protected at the primary amine. Alpha amino acids (2-3 mmoles) were suspended in phosphate buffer (200mM, pH=10) and kept on ice. Fmoc-succinimidal ester [34] was freshly prepared in DMF on ice and slowly added (1mmol was prepared fresh each time and added at three different time points (0, 3, 8 hours) over the 24 hours). The pH was maintained to a range of 9-10 by incremental addition of phosphate buffer (500mM). Reaction mixture was washed 2x100ml Hexane, acidified Conc. HCl to pH 2.5. The resulting precipitant was extracted 3x100ml ethyl acetate, and washed with acidified brine (pH 2.5) 3x300ml, and dilute HCl (pH 2.5) 3x300ml, and dried by rotory evaporation. The resulting solid or oil was triturated with hexanes and recrystallized from ethyl acetate and hexane. Compounds **28**, **30**, **32** were further purified by silca gel chromatography (40/3/2, DCM/MeOH/AcOH) Purity was measured by TLC (40/3/2, DCM/MeOH/AcOH) and with detection by UV and ninhydrin staining. All compounds were analyzed by *1D ^1H NMR (400MHz)*. All the spectra showed the characteristic shift of the HCα proton peak associated with the Fmoc protection, except those with Cα branched residues, and showed the proper proton integration for the Fmoc protection.

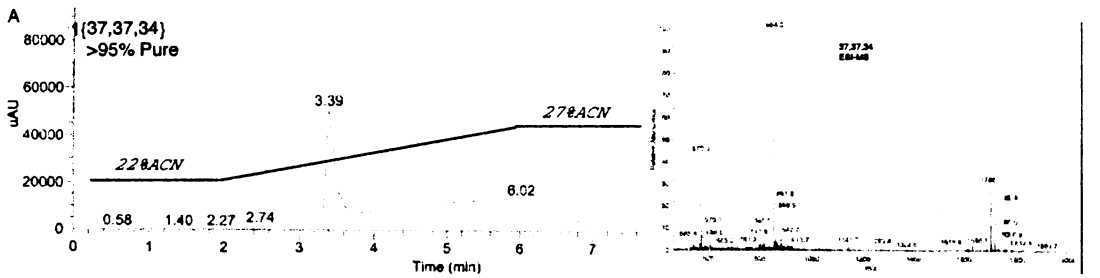


Figure 4-4. Reverse Phase LC/MS Analysis of 1{37,37,34} High Purity

Compounds were analyzed on a Finnegan LCQ. Purified compounds were arrayed in a 96 well plate suspended in 25%ACN/H_2O and chilled to 4°C. 10µl of the each compound were injected onto an Xterra 0.1x50mm RP_{18} column with a flow rate of 200µl/min at 20°C and eluted using a gradient of 22-27% ACN, analyzed by UV absorbance at 215nm and 280nm and ESI-MS. A) High Purity (>95%) Compound **1{37,37,34}**. The full ESI-MS shows the parent, single, and triple charged species of the identified compound, calculated mass (M+H) for 1{37,37,34} = 1705.95, with some impurities shown. The trailing peak between 4-6minures was consistent with a blank (see. **Figure 3**) and all samples; no mass higher than 150 daltons registered during this period on the TIC. Compounds were also tested in MeOH/ H_2O and at a column temperature of 50°C with no observed differences in purity between the compounds. Purity of the compounds was determined as a ratio of the UV Abs 215nm of each peak or the total ion count (TIC) for each chromatogram using whichever gave the lower value.

Synthesis of Labeled Probe

A solution of the linear SRC2-2 (**37,37,37**) (0.8 mmoles) in aqueous DMF (100 µL 1:1, 200 mM sodium phosphate (pH 7): DMF) was treated with Oregon Green 488 succinimidyl ester (Molecular probes 0-6147, 3.2μM in 50µL DMF) and allowed to stir at room temperature for 2 h. The resulting fluorescently labeled peptide **OG-SRC2-2** was purified by RP-HPLC[81].

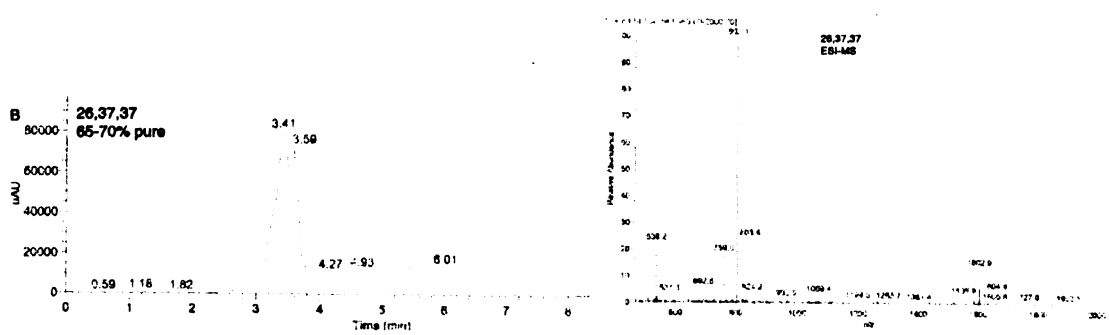


Figure 4-5. Reverse Phase LC/MS Analysis of 26,37,37 (Low Purity)

Purity of the compounds was determined as a ratio of the integrations of the UV Abs at 215nm or the total ion count (TIC) of each peak in the chromatogram at 3.41 minutes, with the target compound, and 3.59 minutes with an impurity, Background baseline from the blank injection (Figure 3) was subtracted from the calculations. The Relative abundance from the low resolution ESI-MS shows the peak identification of the complete chromatogram for 1{26,37,37}. The full ESI-MS shows the parent, single and triple charged species of the identified compound, calculated mass (M+H) for 1{26,37,37} = 1602.89, with some impurities shown. The trailing peak after four minutes registered no mass higher than 150 daltons.

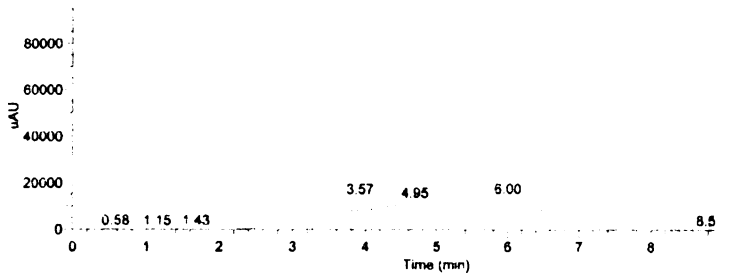


Figure 4-6. Reverse Phase LC/MS Analysis of Blank Injection

A blank injection shows the trailing peak between 2.5-6 minutes was consistent accross all samples with no mass higher than 150 daltons registering during this period on the TIC.

Figure 4-7. Lead Compound 1{37,37,37} Competition Curves

In vitro competition curves are shown for the lead compound **1{37,37,37}** (constrained SRC2-2). Polarization values are plotted on the Y-axis and the $\log_{10}$ of the inhibitor concentration are plotted on the X-axis. The IC_{50} values are the mean values and the 95% confidence interval range of six experiments, triplicate experiments on two different days with two different batches of protein, where inhibitor is evaluated for its ability to displace native SRC2-2 peptide from T_3•hTRβ (λ), E_2•hERα (o) and E_2•hERβ (ν) as assessed using fluorescence polarization (FP) equilibrium competition assays. IC_{50} values are determined as the inhibitor concentration at 50 percent of inhibition as identified by fitting data to one site competition equation y=bottom+(top-bottom)/(1+x/IC_{50}). An IC_{50} value of >100μM indicates the IC_{50} value was determined to be greater than the limits of our assay at 100μM. The difference between each curve was evaluated by the t-test with the following results: T_3•hTRβ vs. E_2•hERα ($p < .001$); T_3•hTRβ vs. E_2•hERβ ($p < .001$); E_2•hERβ vs. E_2•hERα (p= 0.083). The fold difference in IC_{50} results for each NR were as follows: T_3•hTRβ vs. E_2•hERα = 4.7; T_3•hTRβ vs. E_2•hERβ = 3.7;E_2•hERβ vs. E_2•hERα = 1.4. While the control compound was significantly different between each NR it was not selective between the NR.

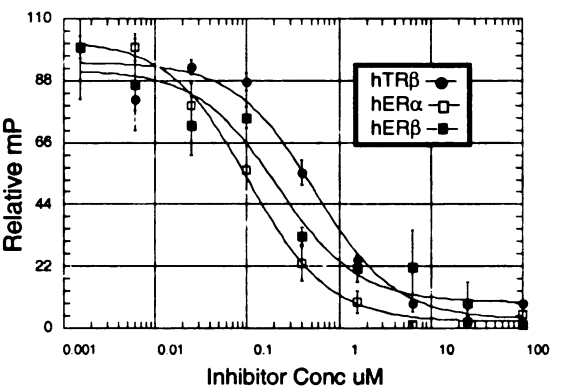
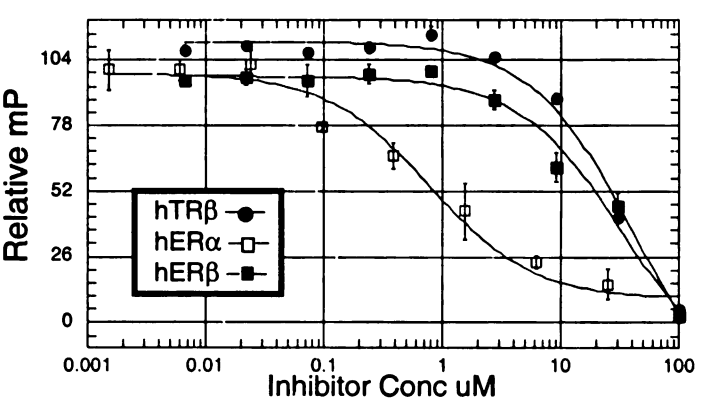


Figure 4-8. 1{37,37,34} Competition Curves

In vitro competition curves are shown for the lead compound **1{37,37,34}**. Polarization values are plotted on the Y-axis and the $\log_{10}$ of the inhibitor concentration are plotted on the X-axis. The IC_{50} values are the mean values and the 95% confidence interval range of six experiments, triplicate experiments on two different days with two different batches of protein, where inhibitor **1{37,37,34}** with tryptophan in the L_3 position is evaluated for its ability to displace native SRC2-2 peptide from T_3•hTRβ (λ), E_2•hERα (o) and E_2•hERβ (ν) as assessed using fluorescence polarization (FP) equilibrium



competition assays. IC_{50} values are determined as the inhibitor concentration at 50 percent of inhibition as identified by fitting data to one site competition equation y=bottom+(top-bottom)/(1+x/IC_{50}). An IC_{50} value of >100μM indicates the IC_{50} value was determined to be greater than the limits of our assay at 100μM.

Pos L1	TRβ T3		ERα E2		ERβ E2	
Comp	IC50	(95% CI)	IC50	(95% CI)	IC50	(95% CI
1	27.7	(14.9-51.3)	0.256	(0.159-0.412)	2.43	(1.33-4.43)
2	22.5	(10.3-49.2)	18.4	(1.69-199)	7.82	(5.27-11.6)
3	7.02	(4.31-11.4)	>100	(0)	20.4	(14.1-29.7)
4						
5	22.1	(14.2-34.5)	1.89	(1.288-2.77)	5.58	(3.97-7.85)
6	54.7	(26.2-114)	0.732	(.5191-1.033)	16.2	(11.6-22.6)
7	1.97	(1.49-2.59)				
8			2.12	(1.394-3.227)	14.9	(9.75-22.6)
9	16.8	(10.5-27.0)				
10	>100	(0)	1.83	(1.33-2.507)	18.6	(12.2-28.3)
11						
12						
13	68.8	26.8-176	3.91	(2.626-5.835)	4.61	(1.21-17.7)
14	5.97	(3.50-10.2)	4.09	(2.637-6.352)	8.76	(5.31-14.4)
15						
16	5.23	(3.03-9.03)	3.59	(2.581-4.993)	16.1	(8.99-28.9)
17	>100		10.6	(5.723-16.64)	42.5	(7.39-245)
18	4.55	(2.55-8.09)				
19						
20	1.43	(.923-2.21)	0.072	(0.037-0.141)	0.118	(0.064-0.215)
21						
22						
23	3.61	(2.32-5.61)	0.144	(0.093-0.221)	0.900	(0.706-1.13)
24	60.3	(18.4-197)	1.47	(.9691-2.238)	9.17	(5.95-14.1)
25	54.6	(20.4-146^)	1.65	(1.106-2.459)	3.36	(2.35-4.81)
26						
27	>100		>100		>100	
28	36.6	(17.6-76.3)	16.0	(8.425-30.4)	21.0	(10.7-41.4)
29	5.98	3.42-10.5	2.12	(1.588-2.817)	4.01	(1.62-9.93)
30						
31	>100		>100		>100	
32	>100		5.26	(3.897-7.097)	>100	
33						
34	40.3	(15.6-104)	0.144	(0.093-0.221)	19.2	(10.9-33.9)
35	>100					
36	9.88	5.38-15.3				
37'						
Constrained SRC2-2	0.67	(.369-1.22)	0.143	(0.111-0.184)	0.180	(0.146-0.221)
37' Free NH_2						
Constrained SRC2-2	0.60	(.371-1.10)	0.161	(0.122-0.214)	0.213	(0.170-0.267)
37'						
Linear SRC2-2	8.02	(4.41-14.6)	0.554	(0.202-1.52)	1.02	(0.625-1.68)
Scrambled SRC2-2	>100		>100		>100	

Table 4-2. IC_{50} Values 1{X,L,L}

The results for the individual compounds with substituted non-natural amino acids at position L_1 are presented vertically with each respective IC_{50} values for each nuclear receptor shown

horizontally. The IC_{50} values are the mean values and the 95% confidence interval range of six experiments, triplicate experiments on two different days with two different batches of protein, where inhibitor is evaluated for its ability to displace native SRC2-2 peptide from T_3•hTRβ, E_2•hERα and E_2•hERβ as assessed using fluorescence polarization (FP) equilibrium competition assays. IC_{50} values are determined as the inhibitor concentration at 50 percent of inhibition as identified by fitting data to one site competition equation y=bottom+(top-bottom)/(1+x/IC_{50}). An IC_{50} value of >100μM indicates the IC_{50} value was determined to be greater than the limits of our assay at 100μM.[A]Some competitions were inconsistent due to light scattering at concentrations of inhibitor above 50μM with poor saturation above 50μM.

Pos L2	TRβ T3	ERα E2	ERβ E2
Comp	IC50 (95% CI)	IC50 (95% CI)	IC50 (95% CI)
1	1.47 (1.01-2.15 [c])	0.334 (0.219-0.509)	>100
2	13.1 (7.20-23.8)	14.1 (6.21-31.8)	>100
3	>100	17.2 (8.47-35.0)	>100
4	4.75 (3.33-6.77)	2.05 (1.03-4.08)	7.56 (4.50-12.7)
5	>100	0.540 (0.343-0.850)	2.92 (2.14-3.98)
6	5.61 (3.16-9.95)	2.29 (1.39-3.79)	3.61 (2.39-5.43)
7	6.04 (4.16-8.78)	0.348 (0.222-0.543)	1.05 (0.723-1.53)
8	0.938 (0.488-1.80)	3.25 (2.07-5.11)	4.80 (0.904-25.5)
9	6.98 (4.82-10.1)	0.037 (0.0035-0.387)	0.570 (0.328-0.993)
10	6.39 (4.23-9.66)	0.424 (0.282-0.639)	1.22 (0.918-1.63)
11	6.35 (4.08-9.91)	0.323 (0.211-0.493)	1.72 (1.36-2.17)
12	10.6 (6.40-17.4)	0.339 (0.192-0.596)	0.749 (0.510-1.10)
13	4.18 2.55-6.86		
14	>100	9.26 (2.31-37.1)	0.824 (0.567-1.20)
15			
16	4.60 (2.14-9.89)	0.230 (0.142-0.373)	2.33 (1.67-3.26)
17			
18	8.43 (5.13-13.9)	17.17 (1.48-199)	1.98 (1.38-2.84)
19	4.30 (2.53-7.33)	0.364 (0.239-0.572)	2.06 (1.47-2.89)
20	0.508 (0.26-0.98)	9.27 (5.57-15.4)	>100
21	7.18 (5.09-10.1)	0.264 (0.176-0.398)	2.31 (1.68-3.17)
22	5.67 (3.38-9.52)	6.61 (4.11-10.7)	11.9 (3.96-35.7)
23	5.45 (3.48-8.55)	17.3 (3.82-78.6)	2.60 (1.83-3.70)
24	4.97 (3.12-7.92)	0.155 (0.096-0.251)	0.774 (0.544-1.10)
25	3.53 (1.37-9.09 [a])	0.944 (0.620-1.44)	5.26 (3.48-7.96)
26	>100		
27	8.11 (5.53-11.9)	0.132 (0.085-0.207)	0.334 (0.191-0.584)
28	>100 [b]	28.6 (10.1-80.8)	>100
29			
30			
31			
32			
33	15.0 (7.80-29.0[b])	24.8 (9.44-65.3)	>100
34	25.1 (12.3-51.3)	>100	16.8 (5.09-55.7)
35	13.0 (7.40-23.0[b])	0.218 (0.142-0.335)	1.95 (1.50-2.54)
36			
37' Constrained SRC2-2	0.6 (0.313-1.2)	0.122 (0.066-0.226)	0.157 (0.122-0.204)
37' Free NH_2 Constrained SRC2-2		0.150 (0.094-0.241)	0.140 (0.104-0.190)

37'			
Linear SRC2-2	13.1 (7.54-22.8)	0.400 (0.335-0.477)	1.74 (1.17-2.59)
Scrambled SRC2-2	>100	>100	>100

Table 4-3. IC_{50} Values for L_2 Substitutions. 1{L,X,L} Same as above.

Pos L3	TRβ T3	ERα E2	ERβ E2
Comp	IC50 (95% CI)	IC50 (95% CI)	IC50 (95% CI)
1	9.49 (6.47-13.9)	3.54 (2.21-5.68)	>100
2	5.15 (1.01-26.2)	16.5 (3.59-75.9)	>100 (0.018-5.43)
3	8.59 (4.09-15.2)	3.84 (2.58-5.72)	>100 (0.004-34.2)
4	5.46 (1.82-16.3)	0.238 (0.041-1.41)	0.881 (0.192-4.05)
5	6.71 (2.27-19.8)	0.160 (0.025-1.05)	>100 (0.040-2.48)
6	3.79 (1.12-12.7)	0.279 (0.053-1.47)	1.69 (0.561-5.09)
7	12.3 (4.32-34.8)	3.41 (2.32-5.01)	21.2 (13.3-34.0)
8	9.24 (5.99-14.4)	0.760 (0.504-1.15)	>100 (0.186-4.50)
9	4.51 (3.08-6.56)	1.81 (1.19-2.77)	2.38 (0.998-5.66)
10	>100	>100	>100
11	8.44 (5.70-12.5)	3.16 (2.19-4.57)	14.0 (8.46-23.1)
12	3.43 (2.38-4.95)	0.754 (0.492-1.16)	4.92 (3.68-6.57)
13	50.7 (22.3-114)	7.67 (5.07-11.6)	>100 (0.414-11.6)
14	89.6 (17.7-453)	0.058 (0.001-6.87)	0.334 (0.028-3.95)
15	9.61 (4.49-20.6)	4.45 (2.89-6.86)	17.9 (11.5-27.7)
16			
17	24.7 (11.9-51.3)	8.85 (5.01-15.6)	15.9 (7.42-33.8)
18	11.2 (6.56-19.2)	18.3 (6.57-51.0)	13.3 (5.63-31.2)
19	5.58 (3.28-9.5)	1.28 (0.841-1.94)	>100 (0.093-6.24)
20	2.86 (2.11-3.88)	0.148 (0.086-0.254)	0.368 (0.213-0.63)
21	30.8 (12.9-73.5)	1.04 (0.670-1.62)	7.38 (5.14-10.6)
22	23.4 (12.0-45.8)	1.08 (0.323-3.61)	>100
23	10.2 (4.86-21.4)	2.10 (1.40-3.16)	18.8 (11.6-30.5)
24	12.3 (5.94-25.5)	8.39 (5.83-12.1)	>100 (0.015-5.79)
25	>100	>100 (0)	>100
26	>100	6.65 (4.36-10.2)	>100 (5.75-21.1)
27		0.352 (0.227-0.546)	1.02 (0.331-3.11)
28			
29	2.33 (1.61-3.38)	0.886 (0.603-1.30)	>100
30	4.86 (3.38-6.99)	13.4 (5.04-35.7)	5.83 (4.01-8.49)
31	9.39 (6.18-14.3)	0.569 (0.365-0.888)	2.77 (2.04-3.75)
32		2.70 (1.84-3.95)	>100
33	13.9 7.28-26.5		
34	85.8 (19.1-386)	0.555 (0.359-0.858)	8.67 (4.48-16.8)
35	9.11 (5.79-14.3)	8.16 (5.00-13.3)	40.9 (21.0-79.5)
36	11.4 (6.09-21.2)	0 (0)	0.347 (0.261-0.461)
37'			
Constrained SRC2-2	0.80 (0.515-1.24)	0.133 (0.082-0.216)	0.223 (0.160-0.310)
37' Free NH_2 Constrained SRC2-2	1.99 (1.54-2.57)	0.173 (0.124-0.240)	0.133 (0.106-0.166)
37' Linear SRC2-2	10.3 (5.88-18.2)	0.334 (0.259-0.431)	1.98 (1.47-2.67)
Scrambled SRC2-2	>100	>100	>100

Table 4-4. IC_{50} Values for L_3 Substitutions. 1{X,X,L}Same as above.

H_6-hERα and H_6-hERβ Protein expression and purification

H_6-hERα LBD and H_6-hERβ LBD were expressed in *Escherichia coli* BL21(DE3). Frozen competent BL21(DE3) cells were thawed and transformed with a PET3-ERα-LBD and PET3-ERβ-LBD in $CaCl_2$. Cells were grown up at 14° C, 1 mM IPTG added at OD600 = 0.7, induced 24 hr at 14°). Cells were lysed by freeze-thaw and sonication in 50 mM sodium phosphate (pH 8.0), 0.3 M NaCl, 10% glycerol, 25 mM β-mercaptoethanol, and 0.1 mM PMSF after incubation with 0.1 mg/ml lysozyme (20 min, 0° C). The lysate was cleared (Ti45, 36,000 rpm, 1 hr, 4° C) and the protein purified using the procedure described above, except in sodium phosphate buffer, and eluted with 12-300 mM imidazole gradient. The liganded [(3,3',5-triiodo-L-thyronine (T3; Sigma)] receptor was isolated by HPLC (TSK-phenyl column (TosoHaas, Philadelphia, PA)). Imidazole was removed by NAP gel filtration (Pharmacia) and protein concentration determined by BCA assay (Pierce). All the proteins were purified beyond the bacterial lysis step and immediately prepared for analysis by *in vitro* competition assay within a 24 hour time period. Extensive storage at 4°c and/or freeze thaw cycles tend to give inconsistent results due to an increase in structural and functional heterogeneity of the protein.

H6-hTRβ-LBD Expression and Purification

hTRβ LBD was expressed in *Escherichia coli* DH5α as described in chapter 2.

Acknowledgment.

We would like to dedicate this chapter to the memory of the great philosopher Ken Whelan of Cole Valley. He always had a great question and a great sense of humor. Thank you Ken!

Chapter V Ligand Selective Inhibition of the Interaction of Steroid Receptor Coactivators and Estrogen Receptor Isoforms

Timothy R. Geistlinger.; Andrea A. McReynolds; R. Kiplin Guy, Ligand Selective Inhibition of the Interaction of Steroid Receptor Coactivators and Estrogen Receptor Isoforms, *Chemistry and Biology*;

Introduction

Despite the high degree of conservation in the interfaces of both nuclear receptor and the steroid receptor coactivator families, it is apparent from our recent studies that there are subtle differences in the hydrophobic interface of each NR (E_2•hERα, E_2•hERβ and T_3•hTRβ) and that the unique characteristics of each receptor interface could be utilized as handles for non-natural LXXLL mimetics to selectively inhibit the interaction as described in the previous chapter.[79-83] All of these inhibitors were constrained α-helical SRC2-2 LXXLL proteomimetics with only a single sidechain replacement at one of the three leucine positions. These results imply that the selectivities are derived from within the three leucine recognition pocket and indicate that there may be more subtle differences within the pocket which is permitting the NR specific selection of these inhibitors. They also may suggest that our conventional, static view of this interaction may be limited and requires further investigation into the mechanisms of action behind this interaction. If these interfaces were more dynamic, we wanted to explore the possibility that these interfaces could be allosterically modulated by different against ligands in a way that did not alter the cononical LXXLL interaction but significantly enough to alter the binding of the non-natural sidechains in our library of inhibitors.

Herein we report our analysis of the structural and energetic foundation of this interaction and the examination of the ligand dependent allosteric modulation of ER:SRC2-2 interface by three related ligands - Estrogen (E_2), Diethylstilbesterol (DES), and Genistein (GEN), with the two ER isoforms. The results show that while there are sub-site differences within the pockets between each receptor which may be contributing to this selectivity, there is no strong structure activity relationship for each of the receptors or the inhibitors. *In vitro* competition studies show that these sub-sites are allosterically modulated by the different ligands in a way that permits the differential selection of SRC2-2 proteomimetic inhibitors without significantly altering the binding of the native three leucine, LXXLL NR box, motif when presented from the same inhibitor scaffold. As a result, it is our hypothesis that the binding mode of these inhibitors on the surface of the receptors is different from that of the native crystal structures. We propose

that either the inhibitors are binding in a slightly different orientation and/or that the receptor surfaces are plastic and are changing in response to the different inhibitors.

The effect of the ligand depends primarily on their ability to differentially alter the structure, stability, and bio-availability of the receptor LBD.[18] Why these unique effects take place is not clear. Estrogen, and estrogen-like ligands function by binding with high affinity to the ligand binding domain (LBD) defining the hydrophobic core of the domain. The ligands allosterically define the overall structure creating a conformational change in the LBD to promote or inhibit the binding of coregulator proteins consisting of coactivators and corepressors. Until now the focus of ligand allosteric effects has been on gross structural changes like the relative modulation of the helix 12 position within the LBD and its ability to create or conceal the SRC binding pocket to act as an agonist or an antagonist, respectively. We wanted to dig a little deeper and explore the hypothesis that different agonist ligands, which are known to modulate the ER isoforms receptor to recruit SRC2-2 peptide, may create subtle differences within the LXXLL leucine binding pocket of ER isoforms that are beyond the reach of the natural leucine sidechains yet are significant enough to be identified by the non-natural chemistry we have introduced into our inhibitor library to reach farther into the binding pocket. To explore this idea we chose to examine subtle differences in the leucine binding pockets of hERα and hERβ in the presence of different ligands, including a synthetic estrogen analogue diethylstilbesterol (DES), and the phytoestrogen genistein (Gen). Both ligands have been identified as partial agonists and have been shown to recruit SRC2-2 for both ER isoforms. Previouse reports of the effects on ERα and ERβ interactions with SRC peptides in the presence of these ligands demonstrates that while there were changes in the recruitment of the SRC NIDs [87] there was little change in the recruitment of the individual boxes.[76] ERα and ERβ have been determined to recruit SRC2 in the presence of E_2, DES and Gen with similar dissociation equilibria.[76] But while each of the ligands does not significantly perturb the binding of the native LXXLL NR box motifs, it is possible that the hydrophobic pocket is modulated in such a way to maintain the binding of the three leucines but that other non-natural leucine mimetics would be significantly effected by the ligand dependent allosteric modulation of the pocket.

To examine this hypothesis we decided to test the same, previously reported α-helical constrained SRC2-2 mimetic inhibitor library[83] with hERα and hERβ in the presence of three different ligands, E_2, the carcinogen DES, and the soybean phytoestrogen GEN. This mimetic library utilizes the previously described α-helical scaffold[81] with a particular conformational constraint of the NR box peptides which produces a proteomimetic, **1{37,37,37}**, and related combinatorial library SRC2-2 mimetics that functionally disrupts the interaction of SRC2-2 with several NR. This approach takes advantage of subtle differences within the highly conserved, hydrophobic pocket, "hot spot" between the nuclear receptors and their coactivators.

Results

Screening for Ligand•NR Selective Proteomimetics

Previous reports of the effects on ERα and ERβ interactions with SRC2-2 peptides in the presence of these ligands indicated that there was little change in the recruitment of the individual boxes.[76] While the ligands do not significantly perturb the binding of the native LXXLL NR box motifs our analysis indicated that the hydrophobic pocket was being modulated in such a way that other non-natural leucine mimetics would be significantly effected. To explore this idea, we screened the library against hERα and hERβ in the presence of different ligands, including a synthetic estrogen analogue diethylstilbesterol (DES) and the phytoestrogen genistein (Gen), both partial agonists that have been shown to recruit SRC2-2 to both ER isoforms.

We initially evaluated each receptor isoform for its ability to bind the three different ligands E_2, DES and Gen (Figure 5-1). The results indicated that each isoform was functionally capable of binding the different ligands that the relative differences were in agreement with the litrature [96]. The OG-SRC2-2 probe bound to each receptor in a ligand dependent manner with Kd's of 299 nM, 310 nM, 450 nM, for E_2•hERα, DES•hERα, Gen•hERα, and 370 nM, 450 nM, and 520 nM, for E_2•hERβ, DES•hERβ, and Gen•hERβ, respectively. The lead compound **1{37,37,37}** competed for each of these interactions with SRC2-2 with IC_{50} values of 120 nM, 290 nM, 110 nM for E_2•hERα, DES•hERα, Gen•hERα; and 151 nM, 261 nM, and 350 nM, for E_2•hERβ, DES•hERβ,

and Gen•hERβ, respectively. There is little selectivity between isoforms with the lead compound relative to the natural affinities of SRC2-2 (Figure 5-2).

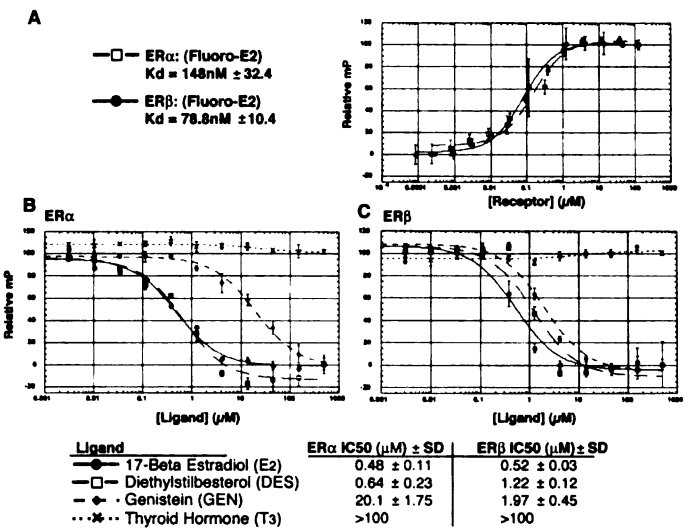
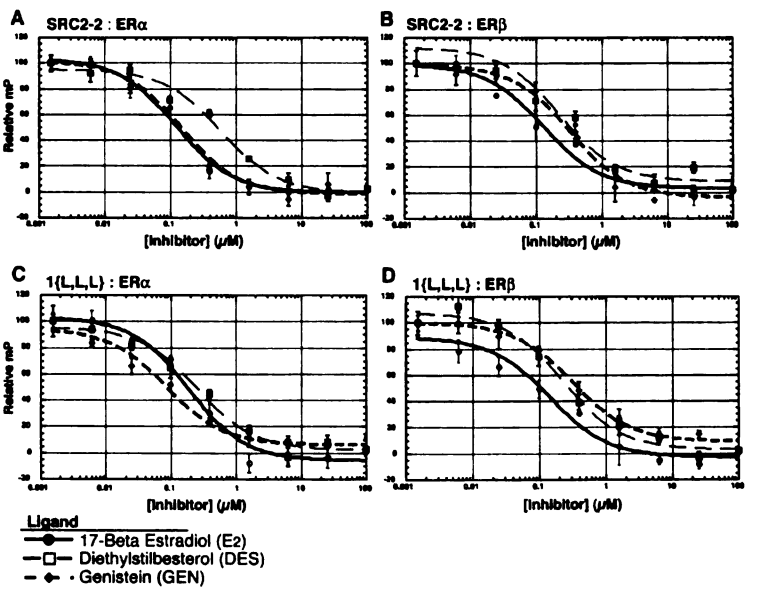


Figure 5-1. ER Isoform Ligand Binding Assay.

Assays were determined by fluorescence polarization competition experiments using a fluorescently labeled estradiol ligand (Fluoro-E_2) (*Invitrogen*). (Panel A) Both ER isoforms were evaluated for their ability to bind a Fluoro-E_2 ligand as well as the ability of each ligand to compete with the fluorscent probe (Panel B and C) (●) 17β-estradiol, (❑) DES, and (◊) Genistein. Ligand binding and competition studies were evaluated in the identical buffer conditions to that of the SRC proteomimetic competition studies.

Figure 5-2. Inhibition of SRC2-2 binding to hERα and hERβ.

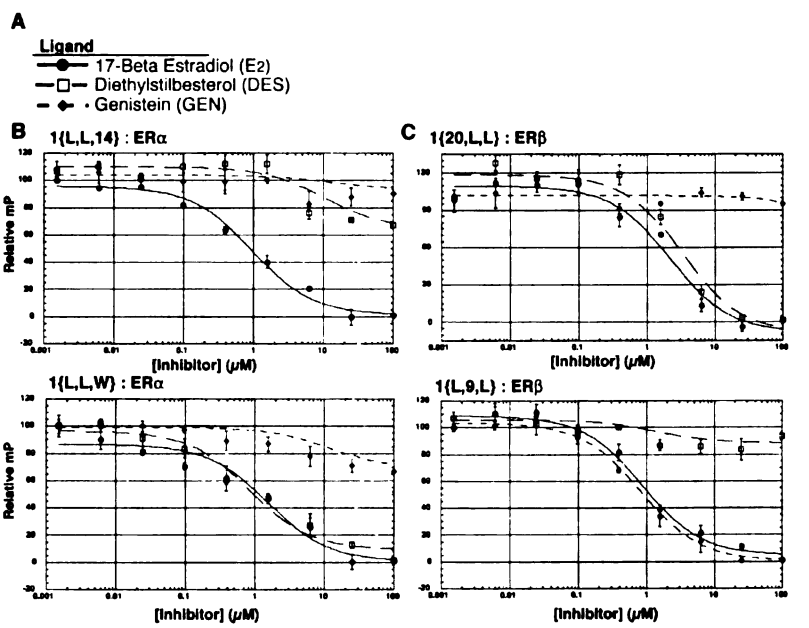
Inhibition of SRC2-2 binding to hERα and hERβ by linear (Panels A,B) and constrained (Panels C,D) SRC2-2 analogues. Assay were determined by fluorescence polarization competition experiments using an Oregon Green 488 labeled SRC2-2 (FITC-SRC2-2) peptide as previously described[81]. Unconstrained **SRC2-2** linear peptide specifically blocks binding of the FITC-SRC2-2 probe with both receptor isoforms (Panel A - ERα and Panel B - ERβ) in the presence of the three ligands (●) 17β-estradiol, (❑) DES, and (◊) Genistein. Likewise, the lead compound **1{L,L,L}** (constrained **SRC2-2**) peptide specifically blocks binding of the FITC-SRC2-2 probe with both receptor isoforms (Panel A - ERα and Panel B - ERβ) in the presence of the three ligands (●) 17β-estradiol, (❑) DES, and (◊) Genistein.



The competitive ability of each of the library members described in Chapter 4 to inhibit the binding of SRC2-2 to the two ER isoforms in the presence of the three ligands was assessed using a fluorescence polarization equilibrium competition assay as

previously described.[79-83] The entire library was then screened with each receptor•ligand pair. Surprisingly, many of the compounds that were previously identified to inhibit the E_2•hERα interaction with SRC2-2 did not inhibit interaction with either one or both of the DES or Gen bound hERα. Likewise, several of the compounds that did not seem to target E_2•hERα inhibited the SRC2-2 interaction with DES•hERα or Gen•hERα (Figure 5-3, 5-4).

Figure 5-3. Selective inhibition of SRC2-2 binding to hERα and hERβ by four of the SRC2-2 proteomimetics.

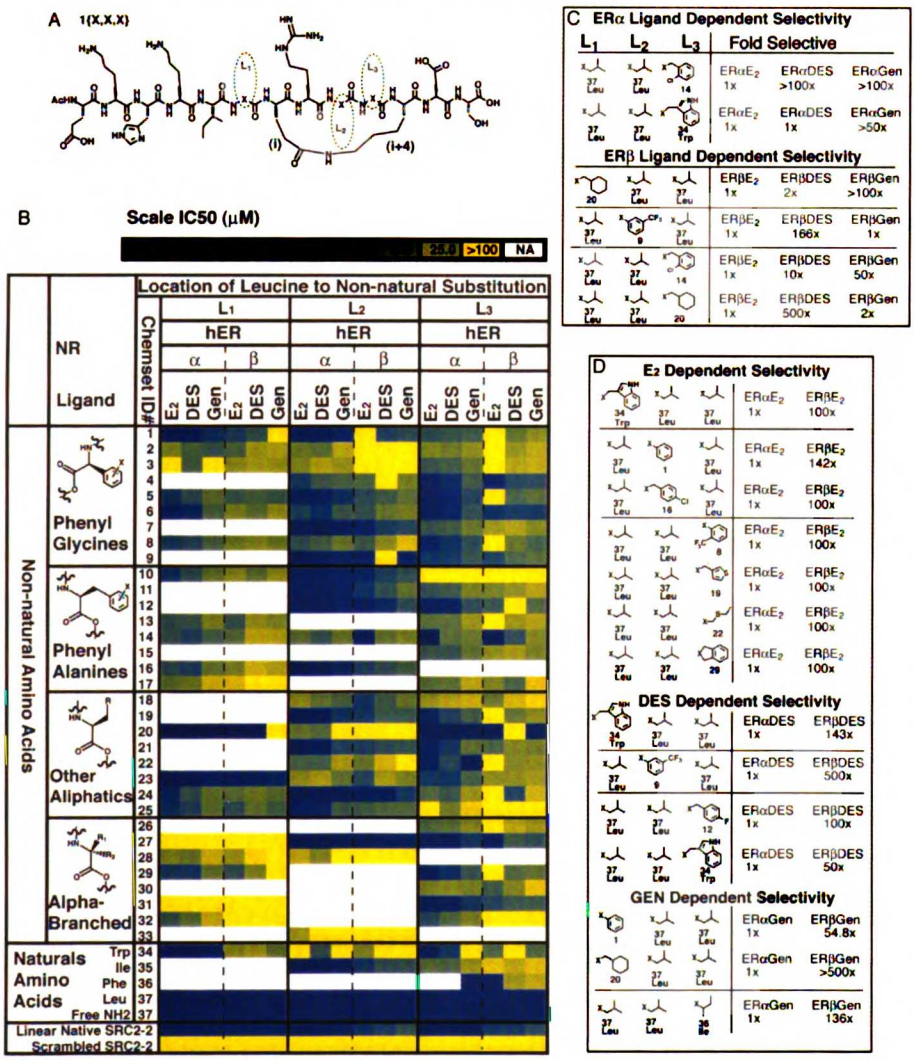


(Panel A) Assay were determined by fluorescence polarization competition experiments using an Oregon Green 488 labeled SRC2-2 (FITC-SRC2-2) peptide as previously described[81]. Each receptor:inhibitor combination was evaluated in the presence of one of the three ligands (Panel A): (●) 17β-estradiol, (❑) DES, and (◊) Genistein. Four compounds are shown as examples of the data represented in Figure 3 f the text: (Panel B) Competition for binding to ERα with inhibitor **1{L,L,14}** and **1{L,L,W}**; and (Panel C) with hERβ and inhibitors **1{20,L,L}** and **1{L,9,L}**.

The selectivity of twelve inhibitors changed by ten fold for hERα in a ligand dependent fashion. At $\mathbf{L_1}$ the *para*-fluorophenylglycine **1{3,37,37}** effectively targeted only the DES•hERα form of the receptor. At $\mathbf{L_2}$, the *meta*-trifluoromethylphenylglycine, **1{37,9,37}**, targeted E_2•hERα, the cyclohexylalanine, **1{37,20,37}**, was selective for both the E_2 and DES liganded forms of hERα, the cyclopentylphenylglycine, **1{37,28,37}** was selective for DES over Gen but not between DES and E_2, and tryptophan, **1{37,34,37}** is highly selective for the DES bound form of hERα. At the $\mathbf{L_3}$ position, three phenylglycine compounds were selective for E_2, 2,3-Difluorophenylglycine, 3,4-Difluorophenylglycine, and 1,5-Difluorophenylglycine, **1{37,37,4}** , **1{37,37,5}**, **1{37,37,6}**. The *ortho*-fluorophenylalanine, **1{37,37,14}** and tryptophan **1{37,37,34}** selected for the DES•ERα ligand.

Figure 5-4. Estrogen Receptor Isoform Ligand Dependent Inhibitory Selectivity Profile of the SRC2-2 Proteomimetics.

Panel A) The structure of the proteomimetic library scaffold **1{X,X,X}** with diversity positions indicated by L_1, L_2, and L_3. 1{37,37,37}: L1 = L2 = L3 = LEU. Panel B) The relative equilibrium 50% inhibitory concentration (IC_{50}) for each library member for competition of the SRC2-2 peptide from each ER isoform (hERα, or hERβ) in the presence of three ligands (E_2, DES, Gen) as determined by two separate quadruplicate *in vitro* fluorescence polarization evaluations is presented in a colorimetric scale that has been normalized to the IC_{50} values of the lead compounds **1{37,37,37}** with gold indicating an IC_{50} >100μM, dark blue indicating an IC_{50} of <100nM, and gradations of color between the two indicating intermediate IC_{50} values. White boxes denote compounds whose synthesis was not achieved. Individual non-natural amino acids are arrayed on the Y-axis, and numbered according to each non-natural sidechain as previously described. The X-axis depicts the position of non-natural amino acid substitution (L_1, L_2, or L_3) and the NR tested. Actual values and the 95% confidence range are presented in Tables 1-3 of the text. Panel C) Selective inhibitors with >50 fold selectivity. The selectivity profile of SRC2-2 proteomimetic inhibitors that are specific to one NR/ligand by >50 fold due to individual side chain substitutions at one Leu position. Selectivity among the tested NR is indicated on the right side with the fold decrease in the IC_{50} relative to the IC_{50} against the receptor for which the compound is most selective. Panel C) Structures of six library members that change their selectivity for hERα and hERβ in a ligand dependent fashion (hERα and hERβ were tested in the presence of estrogen (E_2), the SERM diethylstilbesterol (DES), and the phytoestrogen genistien (GEN). Panel D) The structures of 14 inhibitors that were selective for hERα over hERβ with each ligand.



The specificity of sixteen inhibitors changed for hERβ between the three ligands. At the first leucine position L_1, the phenylglycine ,**1{1,37,37}** was selective for DES and Gen

over E_2. The cyclohexylalanine, **1{20,37,37}** and cyclopentylphenylalanine, **1{29,37,37}** were selective for both E_2 and DES over Gen. At position L_2, phenyl glycine, **1{37,1,37}** was selective for DES and Gen over E_2, while **1{37,4,37}**, and **1{37,5,37}** were selective for both E_2 and Gen over DES, and **1{37,14,37}** was selective for E_2. At position L_3, five phenyl glycines were selective for both DES and Gen over E_2, **1{37,37,1}**, **1{37,37,2}**, **1{37,37,3}**, **1{37,37,5}**, and **1{37,37,8}**; of the phenylglycines, **1{37,37,12}** was selective for E_2 and Gen while **1{37,37,14}** was selective for E_2; **1{37,37,20}** was selective for E_2 and Gen over DES; **1{37,37,24}** was selective for DES; and **1{37,37,29}** was selective for DES and Gen over E_2. Six compounds changed their affinity for hERβ or hERα in a ligand selective manner by >50 fold (Figure 5-4C).

A total of 19 inhibitors were selective for hERα over hERβ by ten fold or more when the receptors were liganded with E_2: four with substitutions at L_1 (**1{6,37,37}**, **1{10,37,37}**, **1{32,37,37}**, **1{34,37,37}**), four at L_2 (**1{37,1,37}**, **1{37,9,37}**, **1{37,16,37}**, **1{37,20,37}**), and 11 at L_3 (**1{37,37,1}**, **1{37,37,3}**, **1{37,37,8}**, **1{37,37,13}**, **1{37,37,19}**, **1{37,37,22}**, **1{37,37,24}**, **1{37,37,26}**, **1{37,37,29}**, **1{37,37,32}**, **1{37,37,34}**. 15 inhibitors were selective for hERα over hERβ when bound to DES, three with substitutions at L_1 (**1{16,37,37}**, **1{32,37,37}**, **1{34,37,37}**), seven at L_2 (**1{37,4,37}**, **1{37,7,37}**,**1{37,9,37}**,**1{37,19,37}**,**1{37,20,37}**,**1{37,34,37}**,**1{37,35,37}**), and five at L_3 **1{37,37,12}**, **1{37,37,23}**, **1{37,37,27}**, **1{37,37,32}**, **1{37,37,34}**). Six were selective for hERα in the presence of Gen, three with substitutions at L_1 (**1{1,37,37}**, **1{20,37,37}**, **1{34,37,37}**, one at L_2 (**1{37,11,37}**, and two at L_3 (**1{37,37,22)**, **1{37,37,36}**). Two of the inhibitors were selective for hERβ, the previously identified *ortho*-fluorophenylalanine at L_2 , **1{37,14,37}**, in the presence of E_2, and tryptophan at L_3, **1{37,37,34}**, in the presence of Gen. Fourteen of these inhibitors were specific to hERα with >50 fold selectivity over hERβ in the presence of one of the three ligands (Figure 5-4D). This set represent the most promising leads for physiologic study or novel therapy. Each of these compounds has the potential to highly selectively abrogate ERα signaling induced by a particular ligand without affecting that of ERβ.

Analysis of Structural Determinants of Selectivity

The interaction modes of a number of the selective inhibitors were evaluated using DOCK (Figure 4A-F). Comparing the results of the *in silico* CombiDOCK screening and

the *in vitro* competition data of the same library reveals that while this method was very useful for enhancing the likelihood of the library producing competent inhibitors, it was not effective for predicting selectivity between the NR. Many of the compounds scored similarly between the receptors and in almost identical minimized positions. The only correlation related to selectivity that is apparent between the *in silico* and *in vitro* binding studies is the size of the sidechains. Larger residues at positions L_1 and L_3 score slightly better when DOCKed with DES:hERα than T_3:hTRβ and that trend is apparent in the competition assays. Several DOCKed inhibitor structures seemed to take advantage of the differences that were identified by the structural and volumetric analysis (Figure 2B-G). Proteomimetics 1{37,8,37} and 1{37,37,14}, for example, bury solvent accessible volume that was identified in ERα and not in TRβ (Figure 4). However, a number of the selective inhibitors would not fit into the pocket unless the Van der Walls constraints were relaxed by 0.2Å (data not shown). These compounds represent selectivity that could not be predicted computationally.

ERα clearly tolerates larger groups within the NR box binding pocket than the crystal structures indicate would be possible. This finding is most likely explained by one of two models: 1) the SRC proteomimetic inhibitors are binding in a different orientation than what is seen from the native peptide crystal structures; and/or 2) the surface of hERα is plastic and is responding to the inhibitors in a way that permits these compounds to unveil new subsites that are energetically favorable. The latter may possible due to the reduced entropic cost of peptide helix formation translating into additional binding energy that perturbs the receptor surface differently than the crystal refined state.

Discussion

Our studies reveal for the first time that the two ER isoforms have a ligand dependent selection for different moieties replacing the leucine side chains when presented on an appropriate proteomimetic scaffold. The three ER ligands (E_2,DES, GEN) are allosterically modulating "sub-sites" of the leucine recognition pockets without perturbing the interactions of the native LXXLL NR box sequences. This suggests that one could potentially simultaneously target one receptor with, for example, an ER

isoform specific ligand and a SRC proteomimetic inhibitor that is selective for that combination.

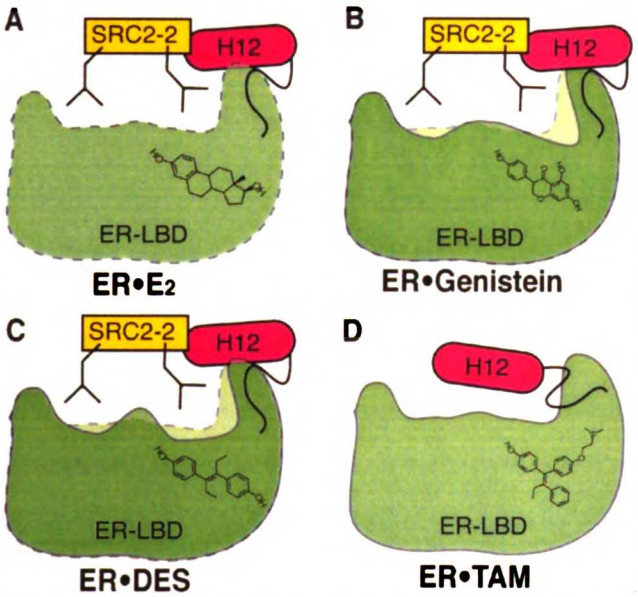


Figure 5-5. Ligand Allosteric Effects Hypothesis

Ligands function to define the globular structure of the LBD. The estrogen receptor (ER-LBD) is known to bind multiple ligands four of which are shown schematically here: A) 17 β- estradiol; B) Genistein; C) Diethylstilbesterol (DES); and Tamoxifen (TAM). Each ligand seems to function differently. Agonist ligands (A-C) position helix 12 (H12) to bind the coativator SRC2 NR box 2 (SRC2-2) with small differences in Kd. Antagoinsts (D) seem to cause H12 to lay over and preclude the binding of the SRC2-2 sequence. Other, more subtle changes may also take place to suggest that these ligands are creating sub-site differences, beyond the "reach" of the canical LXXLL three leucines. These changes may be significant enough to select for different inhibitors with non-natural sidechains. Or these ligands are creating subtle differences in the plasticity of the interface in order to "adapt" to the different inhibitors.

Significance

Ligand dependent nuclear hormone receptor (NR) signaling requires direct interaction between NR and the steroid receptor coactivators (SRC), effected by a series of conserved SRC motifs that are composed of three leucines (NR Box, $L_1XXL_2L_3$). We have previously shown that proteomimetics of the second NR box of SRC2 (SRC2-2) can exploit structural differences between the NR box binding pockets of the thyroid hormone receptor beta (TRβ) bound to thyroid hormone (T_3), and the two estrogen receptor (ER) isoforms (ERα and ERβ), each bound to estrogen (E2). In this report we demonstrate that the same library of SRC2-2 proteomimetic inhibitors can take advantage of differences between the NR box binding pockets of ERα and ERβ in the presence of three different ligands, 17β-estradiol (E_2), Diethylstilbesterol (DES), and Genistein (Gen) to afford specific inhibitors. The selection profile for each ER isoform, and between isoforms, changed depending on which ligand was present. Therefore, the simple hydrophobic SRC binding pocket surface on NR is allosterically modulated differently by each ligand. Fourteen of these inhibitors were specific for preventing recruitment of

SRC2 by hERα with >50 fold selectivity relative to inhibiting the same interaction for hERβ in the presence of one of the three ligands. Each of these compounds has the potential to decouple ERα signaling induced by a particular ligand from that of ERβ. Because this selectivity is induced solely by manipulating the sidechain of one leucine in the analogs, small molecules targeted to this pocket could achieve selectivity for a particular ER•ligand pair by the same mode.

Experimental Procedures

FP ligand binding evaluation.

In 96 well plates, each isoform was serially diluted in binding buffer (50 mM Sodium Phosphate, 100mM NaCl, pH 7.4, 1 mM DTT, 1 mM EDTA, 0.01% NP40, 10% glycerol). Then 10 μl of these diluted protein solutions were added 10μl of the same buffer solution containing 5nM fluorescent labeled estrogen ligand (Fluoro-E_2, *Invitrogen*) in a 384 black, flat bottom plate. The samples were allowed to equilibrate for 30 minutes. Binding was then measured using fluorescence polarization (excitation l 485 nm, emission l at 530 nm) on an Analyst AD (Molecular Devices). Data were analyzed using in Kaleidagraph and the K_d values were obtained by fitting data to the following equation (y=min+(max-min)/1+(x/K_d)^Hillslope) (Figure 5-1A). Polarization values are plotted on the Y-axis and the $\log_{10}$ of the inhibitor concentration are plotted on the X-axis.

FP Ligand Competition.

In 96 well plates, each of the ligands (Estradiol, Diethylstilbesterol, and Genistein) were serially diluted in binding buffer (50 mM Sodium Phosphate, 100mM NaCl, pH 7.4, 1 mM DTT, 1 mM EDTA, 0.01% NP40, 10% glycerol). Then 10 μl of these diluted ligand solutions were added 10μl of the same buffer solution containing 5nM fluorescent labeled estrogen ligand (Fluoro-E_2, *Invitrogen*) and 300nM of each isoform in a 384 black, flat bottom plate. The samples were allowed to equilibrate for 30 minutes. Binding was then measured using fluorescence polarization (excitation l 485 nm, emission l at 530 nm) on an Analyst AD (Molecular Devices). Polarization values are plotted on the Y-axis and the $\log_{10}$ of the inhibitor concentration are plotted on the X-

axis. IC_{50} values are determined as the inhibitor concentration at 50 percent of inhibition as identified by fitting data to one site competition equation y=bottom+(top-bottom)/(1+x/IC_{50}). (Figure 5-1B).

FP SRC2-2 Competition Assyas were performed as described Previously

Pos L1	ERα E2	ERα DES	ERα Gen	ERβ E2	ERβ DES	ERβ gen
Comp	IC50 (95% CI)	IC50 (95% CI)	IC50 (95% CI)	IC50 (95% CI)	IC50 (95% CI)	IC50 (95% CI
1	0.2559 (0.159-0.412)	0.5765 (0.3693-0.9)	1.825 (1.344-2.477)	2.431 (1.334-4.432)	4.886 (3.369-7.087)	100 (
2	18.39 (1.69-199.8)	7.756 (2.755-21.84)	5.321 (3.394-8.343)	7.816 (5.274-11.58)	23.05 (11.92-44.58)	27.2 (10.92-67.63)
3	100 (0)	7.215 (4.936-10.54)	100 (0)	20.44 (14.08-29.67)	23.52 (16.54-33.44)	21.6 (14.98-31.18)
4						
5	1.889 (1.288-2.77)	3.094 (2.076-4.609)	10 (6.781-14.75)	5.583 (3.972-7.847)	7.035 (4.526-10.94)	5.53 (4.068-7.527)
6	0.7321 (0.5191-1.033)	1.423 (1.072-1.889)	6.875 (4.714-10.03)	16.17 (11.55-22.64)	10.54 (6.556-16.96)	19.8 (14.05-28.06)
7						
8	2.121 (1.394-3.227)	2.771 (1.755-4.373)	7.038 (4.806-10.31)	14.85 (9.745-22.62)	17.54 (10.48-29.37)	14.2 (9.036-22.48)
9						
10	1.826 (1.33-2.507)	3.507 (2.506-4.908)	11.62 (7.587-17.81)	18.59 (12.2-28.33)	11.04 (6.462-18.87)	19.2 (12.27-30.03)
11						
12						
13	3.914 (2.626-5.835)	4.569 (2.949-7.079)	14.71 (8.864-24.4)	4.614 (1.206-17.66)	19.6 (10.4-36.92)	17.1 (10.55-27.87)
14	4.093 (2.637-6.352)	2.522 (1.777-3.579)	6.093 (4.253-8.728)	8.76 (5.314-14.44)	27.43 (15.4-48.86)	26.1 (10.3-66.32)
15						
16	3.59 (2.581-4.993)	2.554 (1.706-3.823)	12.1 (7.411-19.75)	16.13 (8.991-28.93)	34.39 (19.11-61.91)	26.7 (13.84-51.64)
17	10.58 (6.723-16.64)	14.14 (9.05-22.09)	36.21 (17.56-74.68)	42.53 (7.391-244.8)	100 (0)	100
18						
19						
20	0.0723 (0.037-0.141)	0.0405 (0.017-0.096)	0.0561 (0.035-0.091)	0.117 (0.0644-0.215)	0.214 (0.159-0.289)	100 (0
21						
22						
23	0.1436 (0.093-0.221)	0.2095 (0.143-0.308)	1.147 (0.840-1.57)	0.895 (0.7062-1.134)	1.233 (0.8485-1.792)	1.51 (1.132-2.03)
24	1.473 (0.9691-2.238)	3.579 (2.34-5.48)	10.76 (6.95-16.7)	9.169 (5.95-14.12)	12.45 (6.942-22.31)	9.13 (4.78-17.5)
25	1.649 (1.106-2.459)	2.58 (1.63-4.08)	9.572 (6.247-14.67)	3.359 (2.35-4.81)	10.22 (6.65-15.7)	6.14 (4.44-8.49)
26						
27	100 (0)	100 (0)	100 (0)	100 (0)	100 (0)	100 (0
28	16.01 (8.425-30.4)	12.54 (7.52-20.9)	33.53 (20.06-56.04)	21 (10.65-41.38)	66.27 (27.6-158)	100 (0
29	2.115 (1.588-2.817)	2.071 (1.394-3.077)	11.44 (8.439-15.51)	4.011 (1.62-9.93)	16.56 (10.0-27.41)	100 (0
30						
31	100 (0)	100 (0)	100 (0)	100 (0)	100 (0)	100 (0
32	5.259 (3.897-7.097)	10.89 (6.621-17.9)	31.24 (16.15-60.43)	100 (0)	100 (0)	100 (0
33						
34	0.1436 (0.093-0.221)	0.2095 (0.142-0.308)	1.147 (0.839-1.568)	19.21 (10.87-33.93)	21.7 (11.32-41.6)	11.7 (6.34-21.9)
35						
36						
37	0.143 (0.111-0.184)	0.233 (0.173-0.314)	0.132 (0.105-0.166)	0.18 (0.146-0.221)	0.311 (0.249-0.388)	0.25 (0.156-0.41)
37-NH2	0.1614 (0.122-0.214)	0.298 (0.149-0.595)	0.312 (0.209-0.465)	0.213 (0.17-0.267)	0.322 (0.125-0.833)	0.42 (0.120-1.463)
Linear 37	0.554 (0.202-1.52)	0.497 (0.199-1.239)	0.465 (0.293-0.737)	1.023 (0.625-1.675)	1.057 (0.517-2.163)	1.22 (0.928-1.609)
Scram	100 (0)	100 (0)	100 (0)	100 (0)	100 (0)	100 (0

Table 5-1. IC_{50} Values Position L_1 Substitution Library 1{X,L,L}.

The results for the individual compounds with substituted non-natural amino acids at position L_1 are presented vertically with each respective IC_{50} values for each nuclear receptor shown horizontally. The IC_{50} values are the mean values and the 95% confidence interval range of six experiments, triplicate experiments on two different days with two different batches of protein, where inhibitor is evaluated for its ability to displace native SRC2-2 peptide from E_2•hERα, E_2•hERβ, DES•hERα, DES•hERβ, Gen•hERα, and Gen•hERβ as assessed using fluorescence polarization (FP) equilibrium competition assays. IC_{50} values are determined as the inhibitor

concentration at 50 percent of inhibition as identified by fitting data to one site competition equation y=bottom+(top-bottom)/(1+x/IC_{50}). An IC_{50} value of >100μM indicates the IC_{50} value was determined to be greater than the limits of our assay at 100μM. ^Some competitions were inconsistent due to light scattering at concentrations of inhibitor above 50μM with poor saturation above 50μM.

Pos L2	ERα E2		ERα DES		ERα Gen		ERβ E2		ERβ DES		ERβ gen	
Comp	IC50	95% CI	IC50	95% CI	IC50	95% CI	IC50	95% CI	IC50	95% CI	IC50	95% CI
1	0.334	0.219-0.509	0.3994	0.2441-0.6535	1.329	0.9699-1.821	100	0	2.61	1.79-3.79	3.661	2.629-5.1)
2	14.1	6.21-31.8	11.24	4.402-28.68	29.36	4.73-182.2	100	0	100	0	100	0
3	17.2	8.47-35.0	24.68	9.42-64.64	20.69	6.925-61.79	100	0	100	0	100	0
4	2.05	1.03-4.08	4.332	2.68-7.003	6.015	3.744-9.663	7.56	4.50-12.7	100	0	15.6	9.548-25.49)
5	0.54	0.343-0.850	0.5738	0.3964-0.8306	1.659	1.155-2.383	2.92	2.14-3.98	5.25	3.31-8.32	9.539	6.485-14.03)
6	2.29	1.39-3.79	1.627	1.075-2.461	3.853	2.485-5.974	3.61	2.39-5.43	10.8 3	6.36-18.5	14.08	8.422-23.53)
7	0.348	0.222-0.543	0.1076	0.06907-0.1676	0.4447	0.2852-0.6933	1.05	0.723-1.53	2.26	1.55-3.17	2.851	1.927-4.217)
8	3.25	2.07-5.11	4.886	3.025-7.89	9.383	5.996-14.68	4.8	0.904-25.5	18.4	7.33-46.5	40.27	19.51-83.11)
9	0.037	0.0035-0.387	0.2294	0.1436-0.3664	0.5982	0.4135-0.8654	0.57	0.328-0.993	10	0	0.6103	0.3227-1.154)
10	0.424	0.282-0.639	0.2779	0.1888-0.4089	0.7535	0.5338-1.064	1.22	0.918-1.63	2.28 2	1.66-3.15	3.383	2.391-4.788)
11	0.323	0.211-0.493	0.3336	0.225-0.4947	0.7676	0.5047-1.167	1.72	1.36-2.17	3.97	2.32-4.13	9.094	6.712-12.32)
12	0.339	0.192-0.596	0.1773	0.1216-0.2585	0.8562	0.6024-1.217	0.749	0.510-1.10	1.5	1.00-2.31	2.875	1.972-4.192)
13												
14	9.26	2.31-37.1	4.144	1.127-15.24	11.42	2.866-45.51	0.824	0.567-1.20	9.94	5.20-19.0	14.21	5.963-33.87)
15												
16	0.23	0.142-0.373	0.424	0.2878-0.6249	1.354	0.9355-1.96	2.33	1.67-3.26	3.17	2.14-4.72	5.667	3.655-8.789)
17												
18	17.17	1.48-199	8.877	1.171-67.27	6.128	2.602-14.43	1.98	1.38-2.84	5.1	3.41-7.70	11.34	6.077-21.15)
19	0.364	0.239-0.572	0.5604	0.3711-0.8461	1.826	1.338-2.491	2.06	1.47-2.89	5.84	3.73-9.15	5.634	3.941-8.054)
20	9.27	5.57-15.4	7.18	3.768-13.68	100	0	100	0	100	0	100	0
21	0.264	0.176-0.398	0.7177	0.4827-1.067	2.448	1.707-3.51	2.31	1.68-3.17	3.7	2.55-5.65	4.724	3.015-7.4)
22	6.61	4.11-10.7	5.969	3.597-9.903	23.18	10.09-53.26	11.9	3.96-35.7	37.2	8.81-158	100	0
23	17.3	3.82-78.6	21.32	2.583-175.9	9.137	4.662-17.91	2.6	1.83-3.70	22.3	5.26-94.7	16.37	5.611-47.75)
24	0.155	0.096-0.251	0.2094	0.1248-0.3512	0.4279	0.2791-0.6559	0.774	0.544-1.10	0.94	0.620-1.44	2.116	1.355-3.304)
25	0.944	0.620-1.44	1.184	0.8423-1.665	4.689	3.372-6.521	5.26	3.48-7.96	7.10	4.41-1.44	7.421	4.927-11.18)
26												
27	0.132	0.085-0.207	0.1251	0.07777-0.2014	0.5044	0.3591-0.7083	0.334	0.191-0.584	0.68	0.509-0.921	1.208	0.8804-1.658)
28	28.6	10.1-80.8	16.12	5.807-44.73	100	0	100	0	100	0	100	0
29												
30												
31												
32												
33	24.8	9.44-65.3	100	0	100	0	100	0	100	0	100	0
34	100	0	3.672	2.193-6.15	100	0	16.8	5.09-55.7	100	0	100	0
35	0.218	0.142-0.335	0.312	0.2251-0.4325	0.8797	0.6451-1.2	1.95	1.50-2.54	3.60	2.43-5.35	3.935	2.799-5.532)
36												
37	0.122	0.066-0.226	0.38	0.257-0.562	0.1253	0.09179-0.171	0.157	0.122-0.204	0.28	0.189-0.434	0.3111	0.2277-0.425)
37-NH2	0.15	0.094-0.241	0.31	0.196-0.490	0.09504	0.06909-0.1307	0.14	0.104-0.190	0.21	0.152-0.308	0.4021	0.2896-0.558)
Linear 37	0.4	0.335-0.477	0.3	0.177-0.509	0.5859	0.4405-0.7791	1.74	1.17-2.59	2.23	1.57-3.18	4.02	2.973-5.428)
Scram	100	0	100	0	100	0	100	0	100	0	100	0

Table 5-2. IC_{50} Values Position L_2 Substitution Library 1{L,X,L}.

Individual IC_{50} values for each inhibitor is listed for the receptor type and conditions as stated in table1 above.

Pos L3	ERα E2	ERα DES	ERα Gen	ERβ E2	ERβ DES	ERβ gen
Comp	IC50 (95% CI)	IC50 (95% CI)	IC50 (95% CI)	IC50 (95% CI)	IC50 (95% CI)	IC50 (95% CI
1	3.54 (2.21-5.68)	4.955 (3.06-8.02)	6.238 (1.69-22.9)	100	13.73 (7.95-23.73)	19.47 (10.77-35.17)
2	16.5 (3.59-75.9)	12.07 (5.14-28.3)	18.48 (6.50-52.5)	100	20.61 (12.27-34.61)	27.97 (12.79-61.17)
3	3.84 (2.58-5.72)	4.897 (3.40-7.06)	10.71 (7.20-15.9)	100	25.28 (15.31-41.74)	28.95 (16.62-50.44)
4	0.238 (0.041-1.41)	1.452 (1.10-1.92)	4.193 (2.85-6.18)	0.881 (0.192-4.05)	2.582 (0.859-7.759)	2.856 (0.780-10.46)
5	0.16 (0.025-1.05)	0.9893 (0.673-1.45)	2.404 (1.69-3.41)	100 (0.039-2.48)	6.176 (4.431-8.609)	7.795 (5.438-11.17)
6	0.279 (0.053-1.47)	1.071 (0.738-1.56)	3.194 (2.15-4.75)	1.69 (0.561-5.09)	2.091 (0.6583-6.644)	1.992 (0.5991-6.62)
7	3.41 (2.32-5.01)	4.842 (3.30-7.11)	11.07 (7.46-16.4)	21.2 (13.3-33.9)	24.68 (15.66-38.9)	16.44 (9.712-27.83)
8	0.76 (0.504-1.15)	1.779 (1.23-2.57)	3.235 (2.34-4.47)	100 (0.186-4.49)	11.77 (7.819-17.71)	6.507 (4.47-9.474)
9	1.81 (1.19-2.77)	2.001 (1.48-2.71)	4.388 (3.07-6.28)	2.38 (0.998-5.66)	7.083 (4.555-11.01)	9.201 (6.34-13.35)
10	100	100	100	100	100	100
11	3.16 (2.19-4.57)	4.243 (2.93-6.15)	13.44 (9.07-19.9)	14 (8.46-23.1)	18.61 (11.34-30.55)	26.12 (16.75-40.72)
12	0.754 (0.492-1.16)	1.19 (0.882-1.61)	3.145 (2.15-4.61)	4.92 (3.68-6.57)	100	8.624 (5.779-12.87)
13	7.67 (5.07-11.6)	7.853 (5.36-11.5)	19.49 (11.7-32.4)	100 (0.414-11.6)	20.5 (9.254-45.41)	25.2 (13.19-48.17)
14	0.058 (0.001-6.87)	7.467 (3.14-17.8)	10.64 (1.77-64.1)	0.334 (0.028-3.95)	3.4 (0.06199-186.4)	16.35 (2.785-95.98)
15	4.45 (2.89-6.86)	3.767 (2.64-5.37)	10.3 (6.98-15.2)	17.9 (11.5-27.7)	22.18 (13.22-37.2)	15.21 (10.05-23.02)
16						
17	8.85 (5.01-15.6)	10.28 (6.59-16.0)	27.68 (12.5-61.4)	15.9 (7.42-33.8)	7.314 (2.281-23.45)	45.74 (13.35-156.7)
18	18.3 (6.57-51.0)	24.89 (4.90-126)	32.16 (10.7-96.5)	13.3 (5.63-31.2)	38.82 (13.29-113.4)	61.75 (3.513-1086.)
19	1.28 (0.841-1.94)	2.191 (1.53-3.14)	4.802 (3.27-7.05)	100 (0.093-6.24)	12.14 (7.69-19.16)	16.15 (11.18-23.34)
20	0.148 (0.086-0.254)	0.1033 (0.066-0.161)	0.4995 (0.362-0.690)	0.368 (0.213-0.636)	100	0.6914 (0.5426-0.88)
21	1.04 (0.670-1.62)	1.882 (1.34-2.65)	8.504 (6.19-11.69)	7.38 (5.14-10.6)	16.85 (10.36-27.41)	14.68 (9.846-21.87)
22	1.08 (0.323-3.61)	2.569 (1.79-3.69)	0.4951 (0.038-8.52)	100	17.44 (11.6-26.21)	19.03 (12.76-28.38)
23	2.1 (1.40-3.16)	1.952 (1.44-2.64)	6.868 (4.79-9.84)	18.8 (11.6-30.5)	19.58 (12.3-31.17)	22.42 (12.64-39.76)
24	8.39 (5.83-12.1)	5.328 (3.54-8.03)	21.19 (12.7-35.4)	100 (0.015-5.79)	2.787 (0.3718-20.9)	33.86 (19.06-60.15)
25	100 (0)	13.88 (6.36-30.3)	57.58 (8.31-398)	100	100	100
26	6.65 (4.36-10.2)	7.818 (5.79-10.6)	16.7 (10.6-26.4)	100 (5.75-21.0)	27.64 (12.01-63.61)	21.32 (10.95-41.54)
27	0.352 (0.227-0.546)	0.3305 (0.223-0.491)	1.94 (1.43-2.63)	1.02 (0.331-3.11)	5.406 (4.009-7.29)	3.59 (2.511-5.133)
28						
29	0.886 (0.603-1.30)	1.617 (1.11-2.36)	4.059 (2.88-5.73)	100	8.115 (5.272-12.49)	11.22 (7.463-16.8)
30	13.4 (5.04-35.7)	14.58 (3.37-63.0)	13.96 (6.15-31.7)	5.83 (4.01-8.49)	16.11 (9.085-28.59)	49.28 (0.7458-325).
31	0.569 (0.365-0.888)	0.9003 (0.626-1.29)	1.99 (1.45-2.73)	2.77 (2.04-3.75)	5.761 (3.882-8.55)	6.252 (4.599-8.49)
32	2.7 (1.84-3.95)	6.874 (4.18-11.3)	15.22 (10.1-22.9)	100	100	100
33						
34	0.555 (0.359-0.858)	2.001 (1.36-2.95)	100	8.67 (4.48-16.8)	100 (0)	8.931 (5.164-15.45)
35	8.16 (5.00-13.3)	6.553 (4.20-10.2)	18.94 (12.1-29.6)	40.9 (20.9-79.5)	29.08 (14.13-59.83)	44.27 (20.79-94.29)
36			0.202 (0.126-0.324)	0.347 (0.261-0.461)	19.63 (2.28-169.)	27.38 (0.9978-751)
37	0.133 (0.082-0.216)	0.29 (0.218-0.386)	0.4642 (0.36-0.599)	0.223 (0.160-0.310)	0.309 (0.191-0.501)	0.229 (0.162-0.32)
37-NH2 Linear	0.173 (0.124-0.240)	0.332 (0.264-0.418)	0.112 (0.069-0.182)	0.133 (0.106-0.166)	0.254 (0.177-0.364)	0.398 (0.339-0.468)
37 Scram	0.334 (0.259-0.431) 100	0.334 (0.211-0.529) 100	0.441 (0.317-0.613) 100	1.98 (1.47-2.67) 100	2.001 (1.549-2.585) 100	3.001 (2.519-3.575) 100

Table 5-3. IC_{50} Values Position L_3 Substitution Library 1{L,L,X}.

Individual IC_{50} values for each inhibitor is listed for the receptor type and conditions as stated in table1 above.

	[Θ]208			[Θ]222		
Comp	**a**	**b**	**c**	**a**	**b**	**C**
1{1}[v]	1.30	-17.87[d]	-0.61	1.30	-11.59[d]	0.49
1{2}[t]	1.52	-0.07	-1.18	2.67	-0.64	-1.14[d]
1{3}[t]	-3.67	-0.85	-3.98	0.23	-0.28	-1.79
1{4}[t]	-	-11.40	-2.84	-	-7.02[d]	-1.48[d]
1{5}[t]	0.29	-	-0.82	0.51	-	0.04

1{6}[t]	2.31	-3.88	-4.59	0.70	-0.81	-2.39
1{7}[t]	-32.73	-8.78	-5.39	-8.97[d]	-6.21[d]	-2.00[d]
1{8}[t]	-	-15.86	-7.74	-	-13.43[d]	-4.27[d]
1{9}[t]	-2.17	-4.11	-3.44	-0.69	-2.22[d]	-1.39[d]
1{10}[u]	-0.84	-4.69	1.10	0.40	-2.43[d]	4.06
1{11}[u]	-	-6.48	-5.29		-3.36[d]	-4.48[d]
1{12}[u]	-	-6.24	-13.05		-2.49[d]	-9.12[d]
1{13}[v]	1.46	-15.15	0.38	2.61	-10.52[d]	0.32
1{14}[u]	-0.62	-0.86	-2.27	0.90	-1.18	-1.52[d]
1{15}[u]	-	-	-3.16	-	-	-3.64[d]
1{16}[u]	-24.45	-5.52		-7.96[d]	-3.10[d]	-
1{17}[u]	-0.40	-	1.78	-2.85	-	0.35
1{18}[w]	-22.41	-6.96	-10.99	-8.65	-4.55[d]	-4.51[d]
1{19}[u]	2.14	-8.63	-4.94	4.20	-5.31	-3.27
1{20}[s]	-16.44	-37.43	-5.61	-6.62[d]	-27.35[d]	-3.94[d]
1{21}[x]		-10.43	-1.88		-8.15[d]	-0.05
1{22}[w]		-4.56	-3.50		-1.11[d]	-1.42[d]
1{23}[w]	-1.23	-4.33	-4.35	0.91	-3.03[d]	-1.74[d]
1{24}[x]	-2.82	-3.93	-10.24	0.25	-3.29	-8.86[d]
1{25}[u]	0.31	-18.86	0.70	-0.11	-20.03[d]	1.26
1{26}[s]		1.18	1.37		0.49	0.59
1{27}[w]	-3.02	-7.20		-1.30[d]	-6.08[d]	
1{28}[y]	-8.23	0.24		0.60	-0.48	
1{29}[y]	-0.20		-5.35	-0.06		-6.20[d]
1{30}[x]			-6.04			-5.66[d]
1{31}[z]	2.52	-4.30	-5.79	2.10	-2.57	-2.84
1{32}[w]	-6.95			1.67		
1{33}[w]		-3.21	-1.36		-1.23[d]	-1.68
1{34}[s]	-4.83	3.00	-4.15	-3.50	0.99	-0.70
1{35}[s]	-11.05	-2.66	-2.11	-4.20	-1.45	-0.49
1{36}[s]	-1.91		-8.89	0.43		-2.99
1{37}[s]	-11.35	-11.01	-4.18	-10.12[d]	-9.71[d]	-2.13[d]
2{37}[s]	1.12	2.62	-0.04	.0.55	0.85	0.11

Table 5-4. Library Circular Dichroism Characterization

Compounds are shown as chemset # and diversity element number. Substituted positions a_x, b_x and c_x are shown in respective columns **a**,**b**,**c**and represent the individual compound data. [d]43 Compounds demonstrated the full CD spectral (200nm - 240nm) double negative elliptical patterns at 208nm and 222nm characteristic of α-helicies and were equivalent to the control **1{37,37,37}**. [e]Compound appeared to break down over a two month time period during freeze thaw cycles in the buffer conditions. All compounds were stored at –80c. [s-z] Indicates vendor in materials section.

Acknowledgements

DOD (predoctoral fellowship of TRG #DAM-17-00-1-0191); the Sidney Kimmel Foundation for Cancer Research, the HHMI Research Resources Program (#76296-549901), the Academic Senate of UCSF, NIH (R01 #DK58080), and the Sandler

Conclusion

Intracellular signaling networks utilize an infinite number of protein-protein interactions to transduce signals throughout the development and maintenance of an organisms life. Aberrant signaling along these pathways has been well correlated with a multitude of disease states. Therefore, inhibiting these interactions successfully would promote not only a better understanding of the role of individual interactions within these signaling pathways, but also engender the design of novel drugs for therapeutic intervention into a multitude of pathologies, both endogenous as well as pathogenic. Until now, however, such efforts have failed primarily due to an inability to achieve selectivity within the large, highly conserved interfaces lacking significant structural handles for drugs to "grab onto". Five years ago we chose to explore novel approaches to target these types of interfaces, to explore the potential role of inhibiting a particular interface, and to do so in a way that had the potential to be expanded to many other types of protein-protein interactions involved in different biological signaling pathways and correlated with disease. We chose the nuclear receptor (NR) steroid receptor coactivator (SRC) interaction (NR•SRC) as a model system because it served not only as an excellent point of intervention into the biology surrounding NR signaling, gene regulation, and related diseases, but also because, like with so many other protein-protein interactions, inhibiting this interaction is plagued with a high degree of structural and phylogenic conservation across both superfamilies of the interacting proteins.

Through structural and energetic evaluations of the NR•SRC interface we hypothesized that there were significant differences in the shape and charge distribution in the conserved, hydrophobic groove of nuclear receptors which could be sufficient enough to allow the development of specific inhibitors. Through the design, synthesis and testing of several scaffolds, the synthesis of a library of 87 novel inhibitors, and *in vitro* competition studies, we identified the first set of inhibitors that selectively inhibited particular nuclear receptor isoforms from coactivator binding in a ligand specific manner. These inhibitors represented the first selective proteomimetics to take advantage of sub-site differences between the leucine binding pockets of E_2•hERα, E_2•hERβ and T_3•hTRβ

when previous studies using genetic selection from random peptide libraries failed to reveal any differences. In the case of the nuclear receptor signaling cascade this methodology revealed not only subtleties of the energetics and sub-site differences within this interface, but also its role in selecting for different coactivators, and the potential to bind sequences that were different from that of the LXXLL motif. As this knowledge evolves, the ability to inhibit these interactions specifically will pave the road to novel development of selective nuclear receptor-regulating chemotherapeutics and more refined therapies for hormone-dependent diseases such as adult onset diabetes, breast, prostate, and skin cancers, cardiovascular disease, inflammatory diseases, and arthritis.

Additionally, there are a multitude of other pathologies not associated with NR signaling but also possessing critical protein-protein interactions that function dynamically with relatively weak affinity for their partners with a similar failure to evolve tight shape complimentarity between receptor and peptide ligand. A particularly tractable subset of such protein interactions associated with various disorders include the SH3 domains,[97] p53•mdm2,[98] Bcl-2,[99] ARPC2 p34•ARPC4 p20,[100] KSHV protease dimerization,[101] CFTR,[102] HSV VP16•TAF_{II},[103] and Gα•Adenylate Cyclase.[104] We expect this methodology to be a powerful approach to clarifying therapeutic targets and generating novel drug leads. This work serves to provide a new framework to not only understand the energetics and subtle differences surrounding protein-protein interactions, but also to identify novel designs for new therapies associated with virtually any disease involving protein interactions.

References

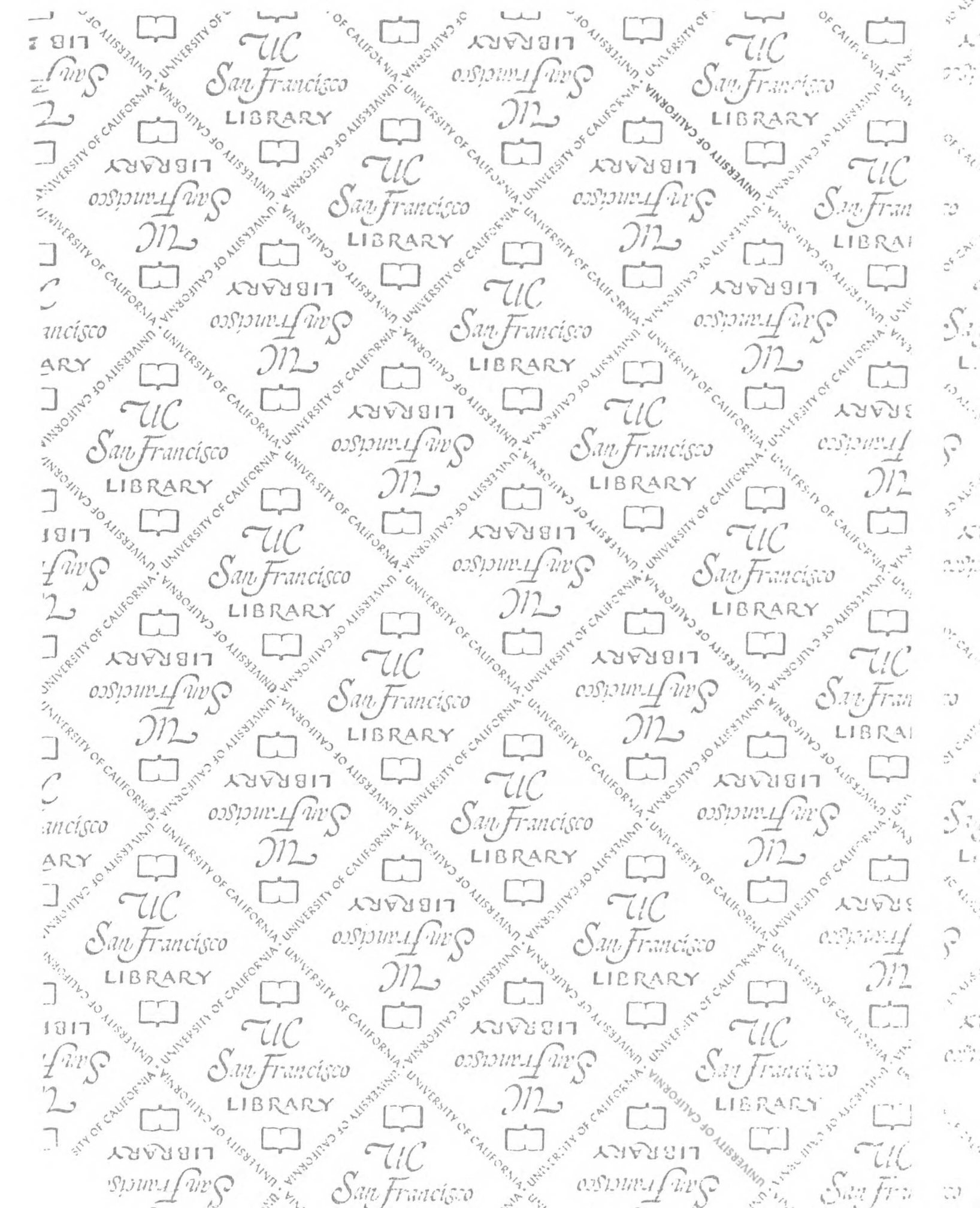
(1) Tsai, M. J.; O'Malley, B. W. *Annu Rev Biochem* **1994**, *63*, 451-486.
(2) McKenna, N. J.; O'Malley, B. W. *Cell* **2002**, *108*, 465-474.
(3) McKenna, N. J.; O'Malley, B. W. *Endocrinology* **2002**, *143*, 2461-2465.
(4) Aranda, A.; Pascual, A. *Physiol Rev* **2001**, *81*, 1269-1304.
(5) Flygare, J. A.; Sutherlin, D. P.; Brown, S. D. *Methods Mol Biol* **2001**, *176*, 353-358.
(6) Shang, Y.; Hu, X.; DiRenzo, J.; Lazar, M. A.; Brown, M. *Cell* **2000**, *103*, 843-852.
(7) Darimont, B. D.; Wagner, R. L.; Apriletti, J. W.; Stallcup, M. R.; Kushner, P. J.; Baxter, J. D.; Fletterick, R. J.; Yamamoto, K. R. *Genes Dev* **1998**, *12*, 3343-3356.
(8) Northrop, J. P.; Nguyen, D.; Piplani, S.; Olivan, S. E.; Kwan, S. T.; Go, N. F.; Hart, C. P.; Schatz, P. J. *Mol Endocrinol* **2000**, *14*, 605-622.
(9) Zhu, Y.; Qi, C.; Jain, S.; Le Beau, M. M.; Espinosa, R., 3rd; Atkins, G. B.; Lazar, M. A.; Yeldandi, A. V.; Rao, M. S.; Reddy, J. K. *Proc Natl Acad Sci U S A* **1999**, *96*, 10848-10853.
(10) Anzick, S. L.; Kononen, J.; Walker, R. L.; Azorsa, D. O.; Tanner, M. M.; Guan, X. Y.; Sauter, G.; Kallioniemi, O. P.; Trent, J. M.; Meltzer, P. S. *Science* **1997**, *277*, 965-968.
(11) Freeman, B. C.; Felts, S. J.; Toft, D. O.; Yamamoto, K. R. *Genes Dev* **2000**, *14*, 422-434.
(12) Freeman, B. C.; Yamamoto, K. R. *Science* **2002**, *296*, 2232-2235.
(13) Feng, X.; Jiang, Y.; Meltzer, P.; Yen, P. M. *Mol Endocrinol* **2000**, *14*, 947-955.
(14) Shao, Z. M.; Shen, Z. Z.; Fontana, J. A.; Barsky, S. H. *Anticancer Res* **2000**, *20*, 2409-2416.
(15) Katzenellenbogen, B. S.; Choi, I.; Delage-Mourroux, R.; Ediger, T. R.; Martini, P. G.; Montano, M.; Sun, J.; Weis, K.; Katzenellenbogen, J. A. *J Steroid Biochem Mol Biol* **2000**, *74*, 279-285.
(16) Dutertre, M.; Smith, C. L. *J Pharmacol Exp Ther* **2000**, *295*, 431-437.
(17) Osborne, C. K.; Zhao, H.; Fuqua, S. A. *J Clin Oncol* **2000**, *18*, 3172-3186.
(18) Shiau, A. K.; Barstad, D.; Loria, P. M.; Cheng, L.; Kushner, P. J.; Agard, D. A.; Greene, G. L. *Cell* **1998**, *95*, 927-937.
(19) Shang, Y.; Brown, M. *Science* **2002**, *295*, 2465-2468.
(20) Kushner, P. J.; Agard, D. A.; Greene, G. L.; Scanlan, T. S.; Shiau, A. K.; Uht, R. M.; Webb, P. *J Steroid Biochem Mol Biol* **2000**, *74*, 311-317.
(21) Shiau, A. K.; Barstad, D.; Radek, J. T.; Meyers, M. J.; Nettles, K. W.; Katzenellenbogen, B. S.; Katzenellenbogen, J. A.; Agard, D. A.; Greene, G. L. *Nat Struct Biol* **2002**, *9*, 359-364.
(22) Wagner, R. L.; Huber, B. R.; Shiau, A. K.; Kelly, A.; Cunha Lima, S. T.; Scanlan, T. S.; Apriletti, J. W.; Baxter, J. D.; West, B. L.; Fletterick, R. J. *Mol Endocrinol* **2001**, *15*, 398-410.

(23) McKenna, N. J.; O'Malley, B. W. *Ann N Y Acad Sci* **2001**, *949*, 3-5.
(24) Moras, D.; Gronemeyer, H. *Curr Opin Cell Biol* **1998**, *10*, 384-391.
(25) Bourguet, W.; Ruff, M.; Chambon, P.; Gronemeyer, H.; Moras, D. *Nature* **1995**, *375*, 377-382.
(26) Bourguet, W.; Ruff, M.; Bonnier, D.; Granger, F.; Boeglin, M.; Chambon, P.; Moras, D.; Gronemeyer, H. *Protein Expr Purif* **1995**, *6*, 604-608.
(27) Renaud, J. P.; Rochel, N.; Ruff, M.; Vivat, V.; Chambon, P.; Gronemeyer, H.; Moras, D. *Nature* **1995**, *378*, 681-689.
(28) Wagner, R. L.; Apriletti, J. W.; McGrath, M. E.; West, B. L.; Baxter, J. D.; Fletterick, R. J. *Nature* **1995**, *378*, 690-697.
(29) Brzozowski, A. M.; Pike, A. C.; Dauter, Z.; Hubbard, R. E.; Bonn, T.; Engstrom, O.; Ohman, L.; Greene, G. L.; Gustafsson, J. A.; Carlquist, M. *Nature* **1997**, *389*, 753-758.
(30) Williams, S. P.; Sigler, P. B. *Nature* **1998**, *393*, 392-396.
(31) Onate, S. A.; Tsai, S. Y.; Tsai, M. J.; O'Malley, B. W. *Science* **1995**, *270*, 1354-1357.
(32) Kamei, Y.; Xu, L.; Heinzel, T.; Torchia, J.; Kurokawa, R.; Gloss, B.; Lin, S. C.; Heyman, R. A.; Rose, D. W.; Glass, C. K.; Rosenfeld, M. G. *Cell* **1996**, *85*, 403-414.
(33) Torchia, J.; Rose, D. W.; Inostroza, J.; Kamei, Y.; Westin, S.; Glass, C. K.; Rosenfeld, M. G. *Nature* **1997**, *387*, 677-684.
(34) Takeshita, A.; Cardona, G. R.; Koibuchi, N.; Suen, C. S.; Chin, W. W. *J Biol Chem* **1997**, *272*, 27629-27634.
(35) Chen, H.; Lin, R. J.; Schiltz, R. L.; Chakravarti, D.; Nash, A.; Nagy, L.; Privalsky, M. L.; Nakatani, Y.; Evans, R. M. *Cell* **1997**, *90*, 569-580.
(36) Voegel, J. J.; Heine, M. J.; Zechel, C.; Chambon, P.; Gronemeyer, H. *Embo J* **1996**, *15*, 3667-3675.
(37) Hong, H.; Kohli, K.; Trivedi, A.; Johnson, D. L.; Stallcup, M. R. *Proc Natl Acad Sci U S A* **1996**, *93*, 4948-4952.
(38) Hong, H.; Kohli, K.; Garabedian, M. J.; Stallcup, M. R. *Mol Cell Biol* **1997**, *17*, 2735-2744.
(39) Glass, C. K.; Rose, D. W.; Rosenfeld, M. G. *Curr Opin Cell Biol* **1997**, *9*, 222-232.
(40) Onate, S. A.; Tsai, S. Y.; Tsai, M. J.; O'Malley, B. W. *Science* **1995**, *270*, 1354-1357.
(41) Torchia, J.; Rose, D. W.; Inostroza, J.; Kamei, Y.; Westin, S.; Glass, C. K.; Rosenfeld, M. G. *Nature* **1997**, *387*, 677-684.
(42) Blanco, J. C.; Minucci, S.; Lu, J.; Yang, X. J.; Walker, K. K.; Chen, H.; Evans, R. M.; Nakatani, Y.; Ozato, K. *Genes Dev* **1998**, *12*, 1638-1651.
(43) Zhu, Y.; Qi, C.; Calandra, C.; Rao, M. S.; Reddy, J. K. *Gene Expr* **1996**, *6*, 185-195.
(44) Xu, J.; Liao, L.; Ning, G.; Yoshida-Komiya, H.; Deng, C.; O'Malley, B. W. *Proc Natl Acad Sci U S A* **2000**, *97*, 6379-6384.
(45) Stallcup, M. R.; Chen, D.; Koh, S. S.; Ma, H.; Lee, Y. H.; Li, H.; Schurter, B. T.; Aswad, D. W. *Biochem Soc Trans* **2000**, *28*, 415-418.
(46) Teyssier, C.; Chen, D.; Stallcup, M. R. *J Biol Chem* **2002**, *277*, 46066-46072.

(47) Hanstein, B.; Eckner, R.; DiRenzo, J.; Halachmi, S.; Liu, H.; Searcy, B.; Kurokawa, R.; Brown, M. *Proc Natl Acad Sci U S A* **1996**, *93*, 11540-11545.
(48) Ma, H.; Baumann, C. T.; Li, H.; Strahl, B. D.; Rice, R.; Jelinek, M. A.; Aswad, D. W.; Allis, C. D.; Hager, G. L.; Stallcup, M. R. *Curr Biol* **2001**, *11*, 1981-1985.
(49) Gehin, M.; Mark, M.; Dennefeld, C.; Dierich, A.; Gronemeyer, H.; Chambon, P. *Mol Cell Biol* **2002**, *22*, 5923-5937.
(50) Xu, J.; Qiu, Y.; DeMayo, F. J.; Tsai, S. Y.; Tsai, M. J.; O'Malley, B. W. *Science* **1998**, *279*, 1922-1925.
(51) Moore, J. and Guy, R.K., Unpublished data in manuscript for 2004.
(52) McInerney, E. M.; Rose, D. W.; Flynn, S. E.; Westin, S.; Mullen, T. M.; Krones, A.; Inostroza, J.; Torchia, J.; Nolte, R. T.; Assa-Munt, N.; Milburn, M. V.; Glass, C. K.; Rosenfeld, M. G. *Genes Dev* **1998**, *12*, 3357-3368.
(53) Ribeiro, R. C.; Apriletti, J. W.; Wagner, R. L.; Feng, W.; Kushner, P. J.; Nilsson, S.; Scanlan, T. S.; West, B. L.; Fletterick, R. J.; Baxter, J. D. *J Steroid Biochem Mol Biol* **1998**, *65*, 133-141.
(54) Ribeiro, R. C.; Apriletti, J. W.; Wagner, R. L.; West, B. L.; Feng, W.; Huber, R.; Kushner, P. J.; Nilsson, S.; Scanlan, T.; Fletterick, R. J.; Schaufele, F.; Baxter, J. D. *Recent Prog Horm Res* **1998**, *53*, 351-392.
(55) Cheskis, B. J.; McKenna, N. J.; Wong, C. W.; Wong, J.; Komm, B.; Lyttle, C. R.; O'Malley, B. W. *J Biol Chem* **2003**.
(56) Chen, D.; Huang, S. M.; Stallcup, M. R. *J Biol Chem* **2000**, *275*, 40810-40816.
(57) Chen, S. L.; Loffler, K. A.; Chen, D.; Stallcup, M. R.; Muscat, G. E. *J Biol Chem* **2002**, *277*, 4324-4333.
(58) Ding, X. F.; Anderson, C. M.; Ma, H.; Hong, H.; Uht, R. M.; Kushner, P. J.; Stallcup, M. R. *Mol Endocrinol* **1998**, *12*, 302-313.
(59) Kalkhoven, E.; Valentine, J. E.; Heery, D. M.; Parker, M. G. *Embo J* **1998**, *17*, 232-243.
(60) Henttu, P. M.; Kalkhoven, E.; Parker, M. G. *Mol Cell Biol* **1997**, *17*, 1832-1839.
(61) Heery, D. M.; Kalkhoven, E.; Hoare, S.; Parker, M. G. *Nature* **1997**, *387*, 733-736.
(62) Yuan, C. X.; Ito, M.; Fondell, J. D.; Fu, Z. Y.; Roeder, R. G. *Proc Natl Acad Sci U S A* **1998**, *95*, 7939-7944.
(63) Chen, S.; Johnson, B. A.; Li, Y.; Aster, S.; McKeever, B.; Mosley, R.; Moller, D. E.; Zhou, G. *J Biol Chem* **2000**, *275*, 3733-3736.
(64) Ko, L.; Cardona, G. R.; Iwasaki, T.; Bramlett, K. S.; Burris, T. P.; Chin, W. W. *Mol Endocrinol* **2002**, *16*, 128-140.
(65) Zhang, H.; Thomsen, J. S.; Johansson, L.; Gustafsson, J. A.; Treuter, E. *J Biol Chem* **2000**, *275*, 39855-39859.
(66) Warnmark, A.; Treuter, E.; Gustafsson, J. A.; Hubbard, R. E.; Brzozowski, A. M.; Pike, A. C. *J Biol Chem* **2002**, *277*, 21862-21868.
(67) Zhang, J.; Lazar, M. A. *Annu Rev Physiol* **2000**, *62*, 439-466.
(68) Webb, P.; Anderson, C. M.; Valentine, C.; Nguyen, P.; Marimuthu, A.; West, B. L.; Baxter, J. D.; Kushner, P. J. *Mol Endocrinol* **2000**, *14*, 1976-1985.
(69) Ghosh, J. C.; Yang, X.; Zhang, A.; Lambert, M. H.; Li, H.; Xu, H. E.; Chen, J. D. *Proc Natl Acad Sci U S A* **2002**, *99*, 5842-5847.

(70) Moraitis, A. N.; Giguere, V.; Thompson, C. C. *Mol Cell Biol* **2002**, *22*, 6831-6841.
(71) Rogatsky, I.; Wang, J. C.; Derynck, M. K.; Nonaka, D. F.; Khodabakhsh, D. B.; Haqq, C. M.; Darimont, B. D.; Garabedian, M. J.; Yamamoto, K. R. *Proc Natl Acad Sci U S A* **2003**, *100*, 13845-13850.
(72) Feng, W.; Ribeiro, R. C.; Wagner, R. L.; Nguyen, H.; Apriletti, J. W.; Fletterick, R. J.; Baxter, J. D.; Kushner, P. J.; West, B. L. *Science* **1998**, *280*, 1747-1749.
(73) Xu, H. E.; Lambert, M. H.; Montana, V. G.; Plunket, K. D.; Moore, L. B.; Collins, J. L.; Oplinger, J. A.; Kliewer, S. A.; Gampe, R. T., Jr.; McKee, D. D.; Moore, J. T.; Willson, T. M. *Proc Natl Acad Sci U S A* **2001**, *98*, 13919-13924.
(74) Xu, W.; Chen, H.; Du, K.; Asahara, H.; Tini, M.; Emerson, B. M.; Montminy, M.; Evans, R. M. *Science* **2001**, *294*, 2507-2511.
(75) Needham, M.; Raines, S.; McPheat, J.; Stacey, C.; Ellston, J.; Hoare, S.; Parker, M. *J Steroid Biochem Mol Biol* **2000**, *72*, 35-46.
(76) Bramlett, K. S.; Wu, Y.; Burris, T. P. *Mol Endocrinol* **2001**, *15*, 909-922.
(77) Chang, C.; Norris, J. D.; Gron, H.; Paige, L. A.; Hamilton, P. T.; Kenan, D. J.; Fowlkes, D.; McDonnell, D. P. *Mol Cell Biol* **1999**, *19*, 8226-8239.
(78) Norris, J. D.; Paige, L. A.; Christensen, D. J.; Chang, C. Y.; Huacani, M. R.; Fan, D.; Hamilton, P. T.; Fowlkes, D. M.; McDonnell, D. P. *Science* **1999**, *285*, 744-746.
(79) Darimont, B. D. *Chem Biol* **2003**, *10*, 675-676.
(80) Geistlinger, T. R., McReynolds, A.C., Guy, R.K. *Chem Biol* **2004**.
(81) Geistlinger, T. R.; Guy, R. K. *J Am Chem Soc* **2001**, *123*, 1525-1526.
(82) Geistlinger, T. R.; Guy, R. K. *J Am Chem Soc* **2003**, *125*, 6852-6853.
(83) Geistlinger, T. R.; Guy, R. K. *Methods Enzymol* **2003**, *364*, 223-246.
(84) COLLABORATIVE COMPUTATIONAL PROJECT, NUMBER 4. 1994. ``The CCP4 Suite: Programs for Protein Crystallography''. Acta Cryst. D50, 760-763.
(85) Generated from an in house script, Volume_Calc, to visuallize and quantify differences in binding pockets. See supplemental materials.
(86) Hanstein, B.; Liu, H.; Yancisin, M. C.; Brown, M. *Mol Endocrinol* **1999**, *13*, 129-137.
(87) Routledge, E. J.; White, R.; Parker, M. G.; Sumpter, J. P. *J Biol Chem* **2000**, *275*, 35986-35993.
(88) UCSF Combidock Kuntz
(89) MDL Information Systems, Inc.
(90) UCSelect Skillman, Thesis 1999
(91) http://www.daylight.com/dayhtml/doc/theory/theory.finger.html
(92) In house perl script
(93) Anchoring Perl script
(94) Perl Script written by A. Shelat, UCSF.
(95) Sun, Y.; Ewing, T. J.; Skillman, A. G.; Kuntz, I. D. *J Comput Aided Mol Des* **1998**, *12*, 597-604.
(96) Invitrogen estrogen receptor ligand competition
(97) Nguyen, J. T.; Porter, M.; Amoui, M.; Miller, W. T.; Zuckermann, R. N.; Lim, W. A. *Chem Biol* **2000**, *7*, 463-473.
(98) Uesugi, M.; Verdine, G. L. *Proc Natl Acad Sci U S A* **1999**, *96*, 14801-14806.

(99) Milella, M.; Estrov, Z.; Kornblau, S. M.; Carter, B. Z.; Konopleva, M.; Tari, A.; Schober, W. D.; Harris, D.; Leysath, C. E.; Lopez-Berestein, G.; Huang, Z.; Andreeff, M. *Blood* **2002**, *99*, 3461-3464.
(100) Robinson, R. C.; Turbedsky, K.; Kaiser, D. A.; Marchand, J. B.; Higgs, H. N.; Choe, S.; Pollard, T. D. *Science* **2001**, *294*, 1679-1684.
(101) Reiling, K. K.; Pray, T. R.; Craik, C. S.; Stroud, R. M. *Biochemistry* **2000**, *39*, 12796-12803.
(102) Naren, A. P.; Cormet-Boyaka, E.; Fu, J.; Villain, M.; Blalock, J. E.; Quick, M. W.; Kirk, K. L. *Science* **1999**, *286*, 544-548.
(103) Hori, R.; Pyo, S.; Carey, M. *Proc Natl Acad Sci U S A* **1995**, *92*, 6047-6051.
(104) Kisselev, O. G.; Meyer, C. K.; Heck, M.; Ernst, O. P.; Hofmann, K. P. *Proc Natl Acad Sci U S A* **1999**, *96*, 4898-4903.


UC
San Francisco
LIBRARY
UNIVERSITY OF CALIFORNIA · UNIVERSITY OF CALIFORNIA