Title
Studies in Structure, Chemistry & Biology Of Receptor-Ligand Interactions for the AMPA & Androgen Receptors

Author
Cruz, Leslie Ann

Publication Date
2011

Peer reviewed|Thesis/dissertation

Studies in Structure, Chemistry & Biology Of Receptor-Ligand Interactions for the AMPA & Androgen Receptors

by

Leslie Ann Cruz

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Chemistry & Chemical Biology

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

In Memoriam

Andrew Braisted (1963-2003)

Warren DeLano (1972-2009)

Two of the best scientists that I had the opportunity work with at

Sunesis Pharmaceuticals.

So much talent. Gone too soon…

Dedication

To my husband, George, and my son, Thomas

My two most favorite men.

To the grandfather I never had:

Dr. David T. Petty, my life-long mentor

To My Family and Friends who have supported and encouraged

me throughout the years:

Mom, Jim, Dad, Mikey, Maria, Lyla

Carolyn, Bob, Melissa, Rob, LeeAnn, Graciela, James

David, Janell, Lori, Becky, Judy, Alex, Gigi

Daniel, Astrid, Dave, Scott, Joice, Kwasi, Marcus

Dan and Monya

Jeanne and Bruce

Acknowledgments

The saying goes, “It takes a village to raise a child”.

It also takes a village to raise a scientist.

Thank You to All My Teachers and Mentors

With Special Thanks To:

St. Benedict's Elementary School

Ilene Hopkins

Memorial Jr. High School

Timothy Sandow

Thornton Fractional South High School

Richard Powell and Ann Rice

The University of Chicago

Viresh Rawal

Argonne National Laboratory

John Hryn

MediChem Research

Raghu Samy, Stuart Feinberg, Shankar Saha, Dimitry Kolton

Sunesis Pharmaceuticals

Andrew Braisted, Dan Erlanson, Jeanne Hardy, Doug Cary, Brian Cunningham, Brian Raimundo, Molly He, Michelle Arkin, Darin Allen, Warren DeLano, Jim Wells

University of California, San Francisco

My Adviser: Robert Fletterick,

My Dissertation Committee: Holly Ingraham, Jack Taunton,

My Orals Committee: Pam England, Jim Wells, Lily Jan, Kevan Shokat

Chris Olson, Charly Craik, Tom Scanlan, Kip Guy, Sue Miller, Dave Agard, Bob Stroud, David Julius, Roger Nicoll, Maia Vinogradova, Fumiaki Yumoto, Phuong Nguyen, Sam Pfaff, Eric Slivka, Jeremey Wilbur, Cindy Benod, Kris Kuchenbecker, Peter Huang, Elena Sablin, Ulrike Boettcher, Kristin Krukenburg, James Kraemer, Mariano Tabios, Rebeca Choy

Collaborators

Marc Cox, Eva Estébanez-Perpiñá, Paul Webb, John Baxter, Stephen Mayo

Preface

My dissertation is comprised of the two projects I worked on during my graduate career. My first project began during my third rotation as a 1st year student in Robert's lab. I had just finished a rotation with Pam England. I wanted to work on a crystallography project before joining her lab. I absolutely fell in love with proteins and x-ray crystallography while working on trying to obtain a crystal structure of ANQX photocrosslinked to the ligand binding domain of the AMPA Receptor. I joined Pam's lab but continued my collaboration with Robert and Eva. The result was a 2008 J. Med. Chem publication. In the process of writing this paper, I switched to Robert's lab. The most major consequence of this switch was that I also had to start over with another project. Robert and I decided on a project that was more involved than just determining an x-ray structure of a protein. Rather, re-engineering the androgen receptor to activate transcription with a novel steroid that does not activate wild-type androgen receptor. My work towards this aim is the second part of my dissertation.

It has been a long hard road to this point in my life, from kindergarten until now. Yet, that road has been full of adventure with starts, stops, bumps, and smooth sailing. Mostly, with bumps though. It has been said that it is not the destination but the journey.

Please Enjoy,

Leslie Ann Cruz

Studies in Structure, Chemistry & Biology Of Ligand-Receptor Interactions for the AMPA & Androgen Receptors

Leslie Ann Cruz

Abstract

Numerous small molecules serve to coordinate many biological activities within cells of multicelluar organisms. These ligand-receptor interactions are vital to the control of cellular processes such as altering cellular membrane potential and activation of gene transcription. This dissertation includes two ligand-receptor interactions projects: validating a small molecule as a selective photoreactive antagonist, ANQX, for AMPA receptors at synapses using x-ray crystallography and mass spectroscopy; and, utilizing a directed evolution yeast based selection assay to re-engineer the androgen receptor to activate transcription with gestrinone, a proof-of-concept drug. These two projects combined provide additional insight into the molecular basis of ligand-receptor interactions of ion channels and steroid nuclear receptors.

Table of Contents

Table of Figures

Chapter 1

Validation of ANQX as a Specific Photoreactive Antagonist of AMPA Receptor

6-Azido-7-nitro-1,4-dihydroquinoxaline-2,3-dione (ANQX) Forms an Irreversible Bond To the Active Site of the GluR2 AMPA Receptor

This article was originally published in Journal of Medicinal Chemistry

Cruz, LA et. al. J. Med. Chem., 2008, 51 (18), pp 5856–5860

DOI: 10.1021/jm701517b

Leslie A. Cruz,‡ Eva Estébanez-Perpiñá,§ Sam Pfaff,∥ Sabine Borngraeber,§ Ning Bao,‡ Justin Blethrow,‡ Robert J. Fletterick,§ and Pamela M. England*⊥#

‡Graduate Program in Chemistry and Chemical Biology, University of California San Francisco. §Department of Biochemistry and Biophysics, University of California San Francisco. ∥Graduate Program in Biophysics, University of California San Francisco. ⊥Department of Pharmaceutical Chemistry, University of California San Francisco. #Department of Cellular and Molecular Pharmacology, University of California San Francisco.

Abstract

AMPA receptors mediate fast excitatory synaptic transmission and are essential for synaptic plasticity. ANQX, a photoreactive AMPA receptor antagonist, is an important biological probe used to irreversibly inactivate AMPA receptors. Here, using X-ray crystallography and mass spectroscopy, we report that ANQX forms two major products in the presence of the GluR2 AMPAR ligand-binding core (S1S2J). Upon photostimulation, ANQX reacts intramolecularly to form FQX or intermolecularly to form a covalent adduct with Glu705.

Introduction

α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) receptors (AMPARs) are a subtype of glutamate-gated ion channels that mediate most fast excitatory synaptic transmission by ensuring rapid responses to synaptically released glutamate.[1–3] In addition, activity-dependent changes in the number of AMPARs at synapses modulates the strength of synaptic transmission, an essential component of the mechanism underlying various forms of synaptic plasticity including learning and memory.[4–8]

AMPARs are composed of four modular subunits (GluR1–4 or GluRA–D), each consisting of an amino-terminal domain (NTD) that modulates receptor assembly, a ligand-binding domain (LBD) that gates the pore of the receptor, three transmembrane segments (M1, M3, M4), a reentrant loop (M2) that lines the pore of the channel, and a

cytoplasmic C-terminal domain that influences receptor trafficking (Figure S1, Supporting Information).[1–3]

High-resolution crystal structures of an engineered ligand-binding core (S1S2J) with several bound ligands have provided insight into the structure and function of full-length receptors.[9,10] Gouaux and co-workers provided the first high-resolution structures of the GluR2 AMPAR ligand-binding core (Figure S1, Supporting Information).[11–13] These structures revealed that the ligand-binding core, formed from two discontinuous polypeptide segments (Figures S1 and S2, Supporting Information), adopts a clamshell-like shape that is open in two states, unliganded (apo) and with a competitive antagonist bound. The clamshell is closed with agonist bound. Notably, structurally related ligands within a given class produce distinct degrees of clamshell closure.[14–17] Coupled with electrophysiological experiments carried out on full-length receptors, these studies suggested that the degree of closure affects the conductance (ion permeation) of the channel, providing a model for channel gating. In addition to modulating channel biophysics, ligand binding also appears to influence the trafficking of AMPARs. For example, both agonists and antagonists have been shown to induce the internalization of AMPARs from neuronal plasma membranes.[18] Although the mechanistic basis for this effect is not understood, it is likely that conformational changes within the ligand-binding domain are translated to the intracellular C-terminal domains, which play a critical role in receptor trafficking.

Quinoxaline-2,3-diones are a major class of competitive AMPAR antagonists, frequently used in studies focused on characterizing the activity of AMPARs.19 Key members of this family of antagonists are 6,7-dinitroquinoxaline-2,3-dione (DNQX) and 6-cyano-7-nitroquinoxaline-2,3-dione (CNQX). Recently we reported the development of ANQX, a new member of the family of quinoxaline-2,3-diones containing an ortho-nitrophenylazide, which irreversibly inhibits AMPARs in the presence of ultraviolet light and provides a means of rapidly inactivating receptors expressed on the surfaces of neurons.[20–22] In the absence of ultraviolet light, ANQX reversibly inhibits AMPARs. The mechanism by which ANQX irreversibly antagonizes AMPARs has not been previously described. Here we report that ANQX forms two products upon photolysis in the presence of the GluR2 ligand-binding core (S1S2J). Following irradiation with ultraviolet light, ANQX loses dinitrogen to form a highly reactive nitrene that either reacts intramolecularly to form [1,2,5]oxadiazolo[3,4-G]quinoxaline-6,7(5H,8H)-dione 1-oxide (FQX) or intermolecularly to form a covalent adduct with Glu705 located in the binding pocket. The high-resolution crystal structure of FQX bound to the S1S2J revealed that FQX binds in the same orientation and promotes the same degree of closure of the S1S2J around the ligand as the previously reported structures complexed with DNQX and CNQX.[12,23] Together, these data indicate a common orientation of quinoxaline-2,3-diones within the AMPAR LBD and reveal the mechanism of receptor photoinactivation with ANQX.

Results

ANQX is a photoreactive quinoxaline-2,3-dione that, upon irradiation with ultraviolet light, loses dinitrogen to form a highly reactive nitrene that undergoes either intra- or intermolecular reactions. We identified the products of both reactions using X-ray crystallography and mass spectrometry. Irradiation of S1S2J-ANQX cocrystals provided a high-resolution structure of the intramolecular reaction product, FQX, reversibly bound to S1S2J. Irradiation of S1S2J in solution in the presence of ANQX yielded sufficient quantities of the intermolecular reaction product to identify GluR705 as the site of cross-linking.

Crystal Structure of the S1S2J–FQX Complex

The high-resolution X-ray structure of the GluR2 ligand-binding core (S1S2J) in complex with FQX was solved at 1.87 Å resolution (Figure 1, Table 1 of Supporting Information). The complex crystallized in the orthorhombic P212121 space group with four protein monomers in the asymmetric unit (Table 1 of Supporting Information). Only two of the monomers are related by noncrystallographic 2-fold symmetry axis. A monomer of the S1S2J dimer complexed with DNQX was used as the search model for molecular replacement. Pairwise superposition of the Cα-atoms of the four molecules yielded pairwise rmsd values ranging between 0.21–0.41 Å.

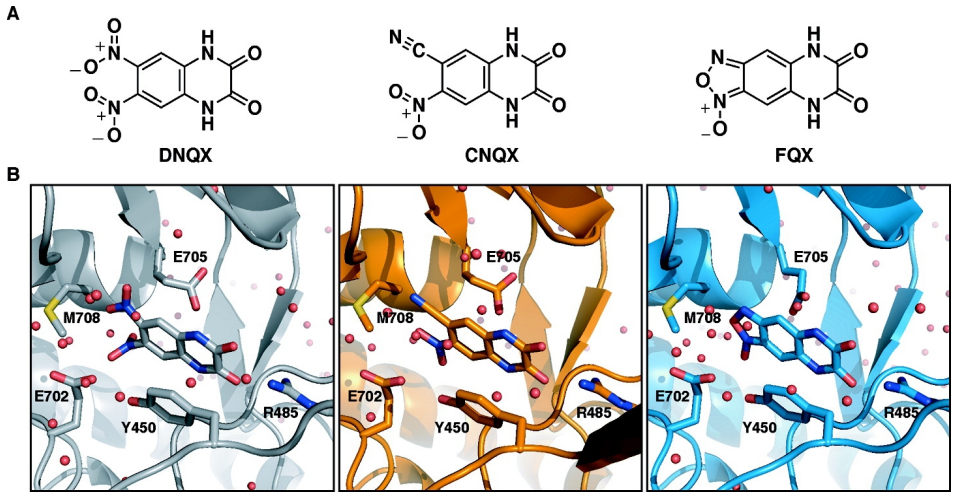


Figure 1

Structural formulas and orientation of the quinoxaline-2,3-diones discussed in the present study within the GluR2 ligand-binding core (S1S2J). (A) The competitive antagonists 6,7-dinitroquinoxaline-2,3-dione (DNQX), 6-cyano-7-nitroquinoxaline-2,3-dione (CNQX), and [1,2,5]oxadiazolo[3,4-G]quinoxaline-6,7(5H,8H)-dione 1-oxide (FQX). (B) Close-up views of DNQX, CNQX and FQX bound to the S1S2J.

Superposition of the S1S2J–FQX monomers with the previously reported S1S2J-DNQX monomer revealed that FQX binds analogously to DNQX within the S1S2J, with an average rmsd of 0.40 Å (Figure 2). The degree of clamshell closure in the S1S2J–FQX structure is identical to that observed in the previously reported S1S2J–DNQX and S1S2J–CNQX structures.[12,23] In all three structures, the clamshell is opened by 15° compared to the glutamate-bound structure. In addition, key residues positioning FQX in the binding site are identical to those orienting DNQX (Figure 2) and CNQX in the ligand-binding core. The quinoxaline ring of FQX is positioned directly below Tyr450, forming a favorable π-stacking interaction. The FQX carbonyl oxygens form hydrogen bonds with Arg485, and the amide nitrogens form hydrogen bonds with the backbone carbonyl of Pro478 and a well-defined water molecule, respectively. The two oxygens of the FQX furoxan ring form a hydrogen bond with a water molecule that is held in position through hydrogen bonds with Tyr405 and Met708. The only binding site residue that differs between the S1S2J–FQX structure and the DNQX and CNQX structures is Glu705, which adopts different conformations in each of the S1S2J–FQX monomer structures due to rotation about the Cβ-Cγ bond.

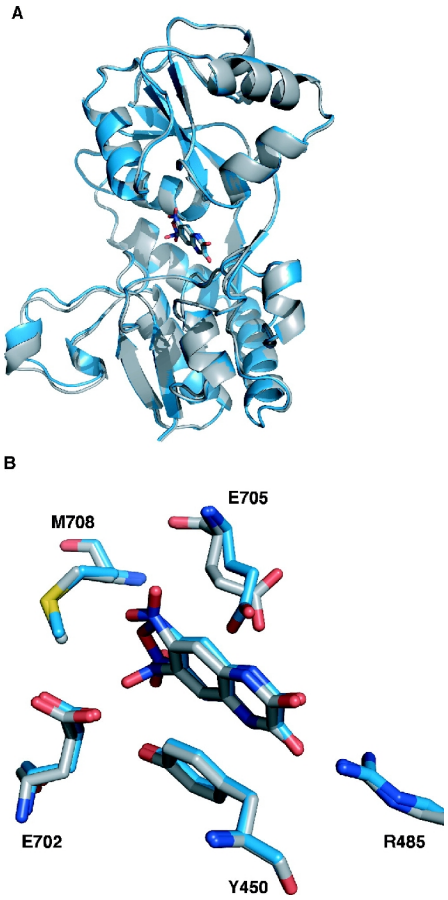


Figure 2

Figure 2. Aligned structures of FQX and DNQX bound to the GluR2 ligand-binding core (S1S2J). (A) FQX (blue ribbon), DNQX (gray ribbon), and CNQX (not shown) produce the same degree of closure of the ligand-binding core. (B) Close-up view of the ligand-binding core showing the residues in close proximity to the ligands FQX and DNQX. Glutamate (E) 705 adopts a different conformation in the FQX structure (blue sticks) than the DNQX structure (gray sticks) and CNQX structure (not shown).

Mass Spectral Analysis of the S1S2J -ANQX Solution Complex

The photocross-linked S1S2J–ANQX adduct was obtained through continuous irradiation and perfusion with "fresh" (unphotolyzed) ANQX over the histidine (6×) tagged S1S2J protein immobilized on Ni-NTA beads. Treatment of S1S2J with ANQX and ultraviolet light, but not ultraviolet light alone, resulted in the formation of covalently modified S1S2J bearing a 220 Da mass increase as determined by electrospray ionization orthogonal-time-of-flight (ESI-o-TOF) (Figure 3A). This is the expected mass increase for the formation of a covalent adduct between photolyzed ANQX (−N2, MW 220) and S1S2J (MW 32218). To determine the location of this adduct, the covalently modified S1S2J was digested with trypsin and the resulting peptides were analyzed by nanoscale liquid chromatography and tandem mass spectrometry (MS/MS). A peptide fragment containing the active site residues Glu705 and Met708 was found bearing a 220 Da adduct in the ANQX-treated sample, which corresponds to the expected mass increase for the formation of a peptide–ANQX fragment (Figure 3B, top panel). In control samples treated with UV light in the absence of ANQX, the 220 Da adduct was not observed (Figure 3B, bottom panel). Comparison of the fragment analysis for the ANQX treated versus untreated samples showed that an adduct formed with Glu705 (y11, E*, top panel, Figure 3B, top panel) and not Met708 (y8, M, Figure 3B, top panel). That is, there is a 220 Da mass difference between the Glu705 fragment in the ANQX treated sample (y11, E*, m/z 1602.72 + NH4, Figure 3B, top panel) compared to that of the Glu705

fragment in the untreated sample (y11, E, m/z 1399.38, Figure 3B, bottom panel). There is no mass difference for the Met708 fragment between the ANQX treated and untreated samples (y8, M, Figure 3B). This peptide fragmentation data provides sufficient coverage to assign the site of covalent modification to Glu705 (E*) and not Met708 (M). We conclude that Glu705 is the preferred site of reaction with photoactivated ANQX.

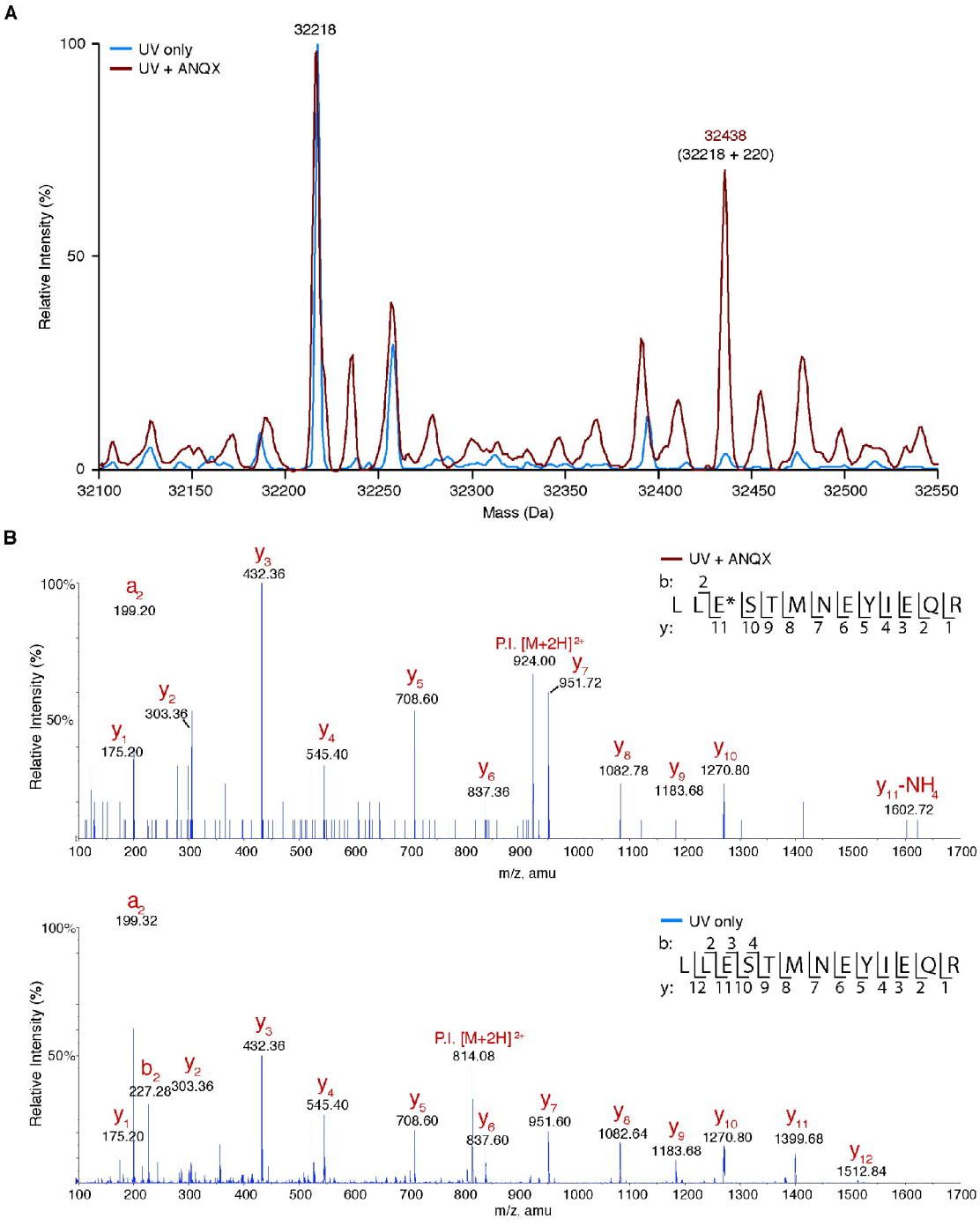
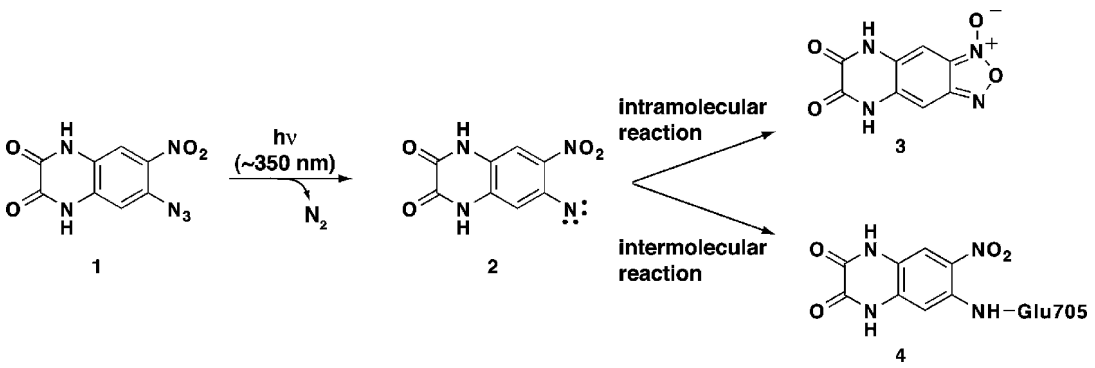


Figure 3. Photocross-linking of ANQX to the GluR2 ligand-binding core (S1S2J) analyzed using mass spectrometry. (A) ESI-o-TOF mass spectrum of the S1S2J

(32218 Da) following irradiation (90 s) in the absence (blue trace) and presence (red trace) of ANQX (100 μM). Irradiation in the presence of ANQX produces a new major peak (32438 Da) that corresponds to the expected increase in mass (+220 Da) for the ANQX−protein adduct. (B) Mass spectra (MS/MS) of the peptide fragment spanning Glu705 following irradiation (90s) in presence (top panel) and absence (bottom panel) of ANQX. Irradiation in presence of ANQX produced a new peak (m/z 1602.72, E*, top panel), corresponding to the expected mass increase for the formation of an ANQX adduct, and the loss of unlabeled Glu705 (m/z 1399.68, E, bottom panel) peak.

Discussion

In the present study, we used X-ray crystallography and mass spectroscopy to identify the products of photolyzing ANQX in the presence of the GluR2 ligand-binding core (S1S2J). Photolysis of ANQX–S1S2J cocrystals yielded a 1.87 Å crystal structure of the intermolecular reaction product, FQX reversibly bound to the GluR2 ligand-binding core (S1S2J). FQX is formed via intramolecular reaction of the nitrene formed following the loss of dinitrogen from ANQX (Scheme 1).[24]



Scheme 1. Photochemistry of ANQX

Upon irradiation with UV light, ANQX (1) forms a highly reactive nitrene (2), which forms two products within the S1S2J binding pocket. The intramolecular reaction results in the formation of FQX (3). The intermolecular reaction results in the formation of a covalent bond with Glu705 in the ligand-binding core (4).Photochemistry of ANQXa

In a previous electrophysiological study, ANQX was used to irreversibly photoinactivate AMPARs on neurons.[21] In these experiments, we noted an initial

increase in the AMPAR current following the first pulse of ultraviolet light, which suggested that a major product of photolyzed ANQX is the formation of a lower affinity antagonist. Consistent with this observation, the IC50 of FQX (7.6 μM[25]) is nearly 8-fold higher than the IC_{50} for ANQX (1.0 μM[20]). The cross-linking efficiency of nitrenes is notoriously poor, with ortho-nitrophenylazides tending to react intramolecularly upon photolysis.[24] Therefore, it is not surprising that we only observed the major reaction product (FQX) reversibly bound to the S1S2J, following irradiation of ANQX–S1S2J cocrystals. An intense effort was made to obtain crystals of the intermolecular reaction product, covalently modified S1S2J–ANQX. However, irradiation of S1S2J–ANQX cocrystallization drops resulted in protein that did not form crystals even after several months and irradiation of the S1S2J immobilized on Ni-NTA beads in the presence of ANQX resulted in protein that was too impure for protein crystallography even after FPLC purification.

Comparison of the FQX structure with the previously reported DNQX and CNQX structures revealed that all three antagonists bind in the same orientation and produce the same degree of clamshell closure within the ligand-binding core. Mass spectral analysis revealed that Glu705 is the site covalently modified by ANQX. On the basis of the position of FQX in the binding pocket in one of the monomers, the nitrene nitrogen on ANQX is positioned 2.5 Å from Glu705. This residue can adopt a range of conformations with distances of 1.7–6.2 Å from the nitrene nitrogen on ANQX. Thus, the cross-link can occur within the range of a single bond without

distortion. Together, these data suggest that the formation of the covalent adduct with photolyzed ANQX does not produce a change in the degree of clamshell closure around the ligand. Because Glu705 is a flexible residue in the S1S2J–FQX crystal structure, it may be the case that only certain rotamers of Glu705 react with the activated ANQX nitrene.

Taken together, these data demonstrate that ANQX is a useful biological probe for inactivating AMPARs because it is not a promiscuous cross-linker. We have shown that ANQX forms a covalent adduct with Glu705 and not Met 708. If this reaction does not occur, ANQX reacts intramolecularly to form FQX, which is a low-affinity antagonist of AMPARs.

Experimental Section

ANQX was synthesized as previously reported.20 The GluR2 ligand-binding core was expressed in Escherichia coli and purified to homogeneity as previously reported by Gouaux and co-workers. The purified S1S2J protein was cocrystallized with ANQX at 4° C using vapor-diffusion (hanging-drop) with a 1:1 ratio of protein to well solution. While submerged in liquid nitrogen, S1S2J–ANQX cocrystals were exposed to ultraviolet light. Synchrotron data for the resulting S1S2J–FQX complex were collected at 100 K on the Advanced Light Source beamline 8.3.1 at Lawrence Berkeley National Laboratory. Data processing was performed using Elves and HKL2000. A complete data set from an S1S2J–FQX crystal diffracted to 1.87 Å,

exhibited the space group P212121 and contained four molecules per asymmetric unit (Table 1 of Supporting Information). Molecular replacement solutions for S1S2J–ANQX crystal structure were obtained using as a search model one of the monomers of S1S2J–DNQX dimer structure. Photocross-linking of ANQX to the S1S2J was accomplished by irradiating a solution of the 6-HIS tagged S1S2J immobilized on Ni-NTA beads in the presence of ANQX. The resulting sample was denatured and digested with trypsin/chymotrypsin and then analyzed by nanoscale LC/MS2 using a Q TRAP mass spectrometer coupled to a liquid chromatography system. Tandem mass spectra were acquired automatically in IDA mode, and the resulting data were analyzed with MASCOT.

Acknowledgments

We thank James Holton at the Advance Light Source (ALS) 8.3.1 beamline, Pascal Egea (UCSF) for assistance collecting and processing data, and Eric Gouaux for the S1S2J construct. This work was supported by grants from the Sandler Foundation and the Herbert Boyer Foundation.

Footnotes

†The structure S1S2J-FQX has been deposited in the PDB and has been assigned the following code: 3BKI.

Abbreviations: AMPAR, α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptor; ANQX, 6-azido-7-nitro-1,4-dihydroquinoxaline-2,3-dione; DNQX, 6,7-

dinitroquinoxaline-2,3-dione; CNQX, 6-cyano-7-nitroquinoxaline-2,3-dione; FQX, [1,2,5]oxadiazolo[3,4-G]quinoxaline-6,7(5H,8H)-dione 1-oxide; S1S2J, GluR2 ligand-binding core; PDB ID: 3BKI.

Supporting Information Available: Experimental methods, a cartoon depicting the portion of the GluR2 subunit corresponding to the S1S2J domain, 2Fo – Fc omit electron density for FQX and surrounding ligands, and statistics for data collection and refinement. This material is available free of charge via the Internet at http://pubs.acs.org.

References

1. Hollmann M, Heinemann S. Cloned glutamate receptors. Annu Rev Neurosci. 1994;17:31–108.
2. Dingledine R, Borges K, Bowie D, Traynelis SF. The glutamate receptor ion channels. Pharmacol Rev. 1999;51:7–61.
3. Palmer CL, Cotton L, Henley JM. The molecular pharmacology and cell biology of alpha-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptors. Pharmacol Rev. 2005;57:253–277.
4. Bliss TVP, Collingridge GL. A Synaptic Model of Memory: Long-Term Potentiation in the Hippocampus. Nature. 1993;361:31–39.
5. Bredt DS, Nicoll RA. AMPA receptor trafficking at excitatory synapses. Neuron. 2003;40:361–379.
6. Song I, Huganir RL. Regulation of AMPA receptors during synaptic plasticity. Trends Neurosci. 2002;25:578–588.
7. Malinow R, Malenka RC. AMPA receptor trafficking and synaptic plasticity. Annu Rev Neurosci. 2002;25:103–126.
8. Collingridge GL, Isaac JR, Wang YT. Receptor trafficking and synaptic plasticity. Nat Rev Neurosci. 2004;5:952–961.
9. Gouaux E. Structure and function of AMPA receptors. J Physiol. 2004;554:249–253. [PMC free article]

10. Mayer ML. Glutamate receptor ion channels. Curr Opin Neurobiol. 2005;15:282–288.

11. Armstrong N, Sun Y, Chen GQ, Gouaux E. Structure of a glutamate-receptor ligand-binding core in complex with kainate. Nature. 1998;395:913–917.

12. Armstrong N, Gouaux E. Mechanisms for activation and antagonism of an AMPA-sensitive glutamate receptor: crystal structures of the GluR2 ligand binding core. Neuron. 2000;28:165–181.

13. Sun Y, Olson L, Horning M, Armstrong N, Mayer M. Mechanism of glutamate receptor desensitization. Nature. 2002;417:245–253.

14. Jin R, Banke TG, Mayer ML, Traynelis SF, Gouaux E. Structural basis for partial agonist action at ionotropic glutamate receptors. Nat Neurosci. 2003;6:803–810.

15. Hogner A, Kastrup JS, Jin R, Liljefors T, Mayer ML. Structural basis for AMPA receptor activation and ligand selectivity: crystal structures of five agonist complexes with the GluR2 ligand-binding core. J Mol Biol. 2002;322:93–109.

16. Jin R, Gouaux E. Probing the function, conformational plasticity, and dimer–dimer contacts of the GluR2 ligand-binding core: studies of 5-substituted willardiines and GluR2 S1S2 in the crystal. Biochemistry. 2003;42:5201–5213.

17. Jin R, Clark S, Weeks AM, Dudman JT, Gouaux E. Mechanism of positive allosteric modulators acting on AMPA receptors. J Neurosci. 2005;25:9027–9036.

18. Lin JW, Ju W, Foster K, Lee SH, Ahmadian G. Distinct molecular mechanisms and divergent endocytotic pathways of AMPA receptors internalization. Nat Neurosci. 2000;3:1282–1290.

19. Nikam SS, Kornberg BE. AMPA receptor antagonists. Curr Med Chem. 2001;8:155–170.

20. Chambers JJ, Gouda H, Young DM, Kuntz ID, England PM. Photochemically knocking out glutamate receptors in vivo. J Am Chem Soc. 2004;126:13886–13887.

21. Adesnik H, Nicoll RA, England PM. Photoinactivation of native AMPA receptors reveals their real-time trafficking. Neuron. 2005;48:977–985.

22. England PM. Rapid photoinactivation of native AMPA receptors using ANQX. Sci STKE. 2006;331:11.

23. Menuz K, Stroud RM, Nicoll RA, Hays FA. Subunits switch AMPA receptor antagonists to partial agonists. Science. 2007;318:815–817.

24. Murata S, Tomioka H. Photochemistry of o-Nitrophenylazide in Matrices. The First Direct Spectroscopic Observation of o-Dinitrosobenzene. Chem Lett. 1992:57–60.

25. Baudy RB, Sulkowski TS. 5h,8h-2-oxa-1,3,5,8-tetraaza-cyclopenta[b]-naphthalene-6,7-diones neuroprotective agents U S. USXXAM US 5719153 A 19980217 CAN 128:176169 AN 1998:146566 1998.

Supporting Information

Experimental

Expression and Purification of the GluR2 S1S2J Domain

The GluR2 ligand-binding core (S1S2J) construct, consisting of the GluR2 S1 segment linked to the GluR2 S2 segment via two amino acids (GT), was kindly provided by Dr. Eric Gouaux (HHMI Investigator, Oregon Health & Science University). The S1 and S2 domains correspond to amino acids 390–506 and 632–763, respectively, in the full-length GluR2 subunit. The S1S2J core was expressed in *E. coli* and purified to homogeneity as previously reported Gouaux and co-workers.[1]

Crystallization, Structure Determination and Refinement

The purified S1S2J protein was co-crystallized with the competitive antagonist ANQX. Apo-S1S2J was concentrated to 10 mg/mL in 10 mM HEPES buffer (10 mM HEPES, 20 mM NaCl, 1 mM EDTA, pH 7.0) and incubated with ANQX (6 mM final concentration). Crystals were grown at 4° C using vapor-diffusion (hanging-drop) with a 1:1 ratio of protein to well solution. Crystals grew to full dimensions after one month in 12.5%-25% PEG 4000, 0.25-0.4 M ammonium sulfate, and 0.1 M NaOAc pH 5.0. Just prior to data collection, the crystals were soaked in 25% ethylene glycol as cryoprotectant and then flash-cooled in liquid nitrogen. While submerged in liquid nitrogen, S1S2J-ANQX cocrystals were exposed to ultraviolet light for 10 s using a Hg/Xe arc lamp (1000 Watt) outfitted with a UV

bandpass filter (#51660, Oriel Instruments: transmitting 300–400 nm, peak 360 nm) and a heat absorbing filter (#51944, Oriel Instruments: transmitting 300- 1000nm). Synchrotron data for S1S2J-FQX was collected at 100K on the Advanced Light Source beamline 8.3.1 at Lawrence Berkeley National Laboratory. Data processing was performed using Elves[2] and HKL2000. For additional details see Table 1 (supplementary material).

A complete dataset from an S1S2J-FQX crystal diffracted to 1.87 Å and exhibits the space group $P2_12_12_1$ and contains four molecules per asymmetric unit (Table 1). Molecular replacement solutions for S1S2J-ANQX crystal structure was obtained using as a search model one of the monomers of S1S2J-DNQX dimer structure.[3] The molecular replacement solutions were obtained using rotation and translation functions from Crystallography & NMR Systems (CNS, http://cns.csb.yale.edu/v1.1/). Protein model refinement consisted of simulated annealing, group and individual B-factor refinement, and conjugate gradient minimization in CNS followed by model building (monitored by free-R factor) using COOT.[4,5] Visual inspection of electron density using COOT allowed identification of the ligand FQX bound to the S1S2J domain. Using CNS, a composite omit map was also calculated in which a different 5% of the model was omitted in an attempt to minimize model bias. Calculation of the electron density maps and crystallographic refinement was performed with CNS using the target parameters of Engh and Huber.[6] Several cycles of model building, conjugate gradient minimization and simulated

annealing using CNS resulted in models with good stereochemistry. A Ramachandran plot shows that all but three of the residues fall into the favored regions. The statistics for data collection and refinement of each one of the data sets are in Table 1.

Photocrosslinking ANQX to S1S2J

A solution of Ni-NTA beads bound to the 6-HIS tagged S1S2J in 10 mM HEPES buffer was placed on a piece of filter paper (Whatmann) in a Petri dish outfitted with a perfusion inlet, supplying unphotolyzed ANQX (100 μM), and an outlet attached to an in house vacuum, removing photolyzed ANQX from the beads. Unphotolyzed ANQX in 10 mM HEPES buffer was continuously perfused (20 mL/min) over the S1S2J Ni-NTA beads while exposing the beads to ultraviolet light for 90 s (1000 Watt Hg/Xe arc lamp outfitted with a UV bandpass filter (#51660,Oriel Instruments: transmitting 300–400 nm, peak 360 nm) and a heat absorbing filter (#51944, Oriel Instruments: transmitting 300-1000nm). In control experiments S1S2J Ni-NTA beads were similarly irradiated in the absence of ANQX. The S1S2J Ni-NTA beads were washed with 10 mM HEPES buffer to remove any remaining photolyzed ANQX. The S1S2J domain was eluted with elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidazole, pH 8.0), dialyzed against 10 mM HEPES, pH 7.0, 20 mM NaCl, 1 mM EDTA, and concentrated (4 mL 10,000 MWCO Amicon centrifuge tubes, Millipore). The samples were analyzed by highthroughput mass spectrometry using a CIT Analytics Autosampler and a LTC premier mass spectrometer (Waters Micromass).

Trpysin-digestion and Tandem Mass Spectrometric Analysis

The experimental (UV light, ANQX) and control (UV light only) S1S2J cores (2.5 g) were treated with 8 M urea and sonicated for 10 min in a 37 °C water. The resulting denatured proteins were digested overnight at 37 °C with a mixture of trypsin/chymotrypsin. Approximately five picomoles of each sample were then analyzed by nanoscale LC/MS^2 using a Q TRAP mass spectrometer (Applied Biosystems, Foster City, USA) coupled to an LC Packings Ultimate/Famos/Switchos liquid chromatography system (Dionex). Peptides were resolved over a 75 micron x 150 mm C18 column using a two hour gradient at a flow rate of 150 nL/min. Tandem mass spectra were acquired automatically in IDA mode, and the resulting data were analysed with MASCOT (Matrix Science).

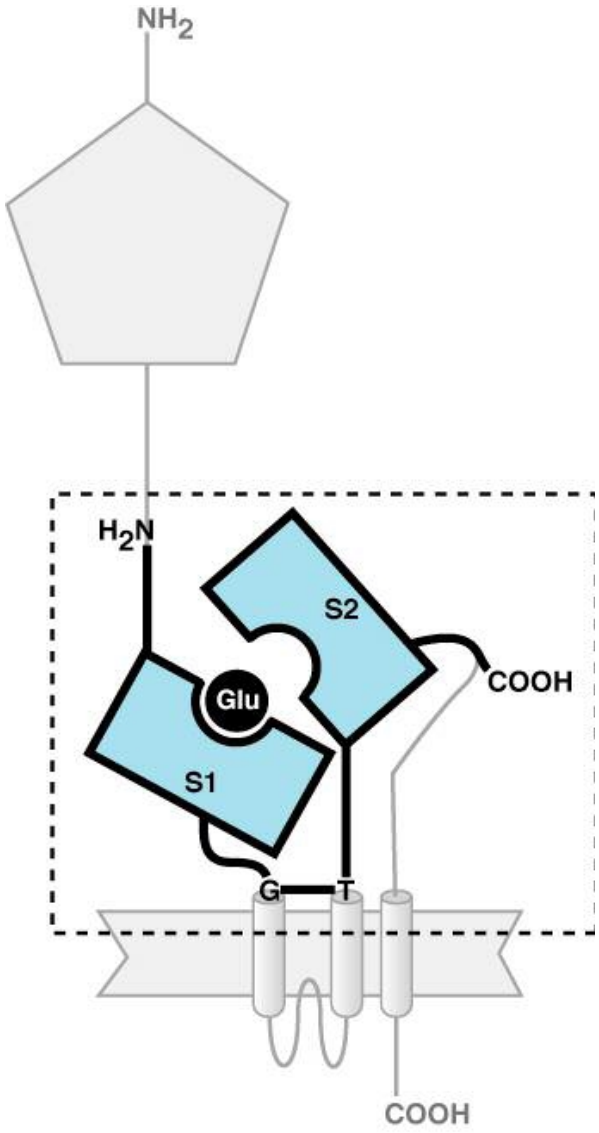


Figure S1. Cartoon depicting the AMPAR ligand-binding core (S1S2J) with respect to the subunit topology. The portion of the subunit corresponding to the S1S2J (dashed box) is comprised of the S1 segment linked to the S2 segment via two amino acids (GT). The S1 and S2 segments correspond to amino acids 390–506 and 632–763, respectively, in the full-length GluR2 subunit.

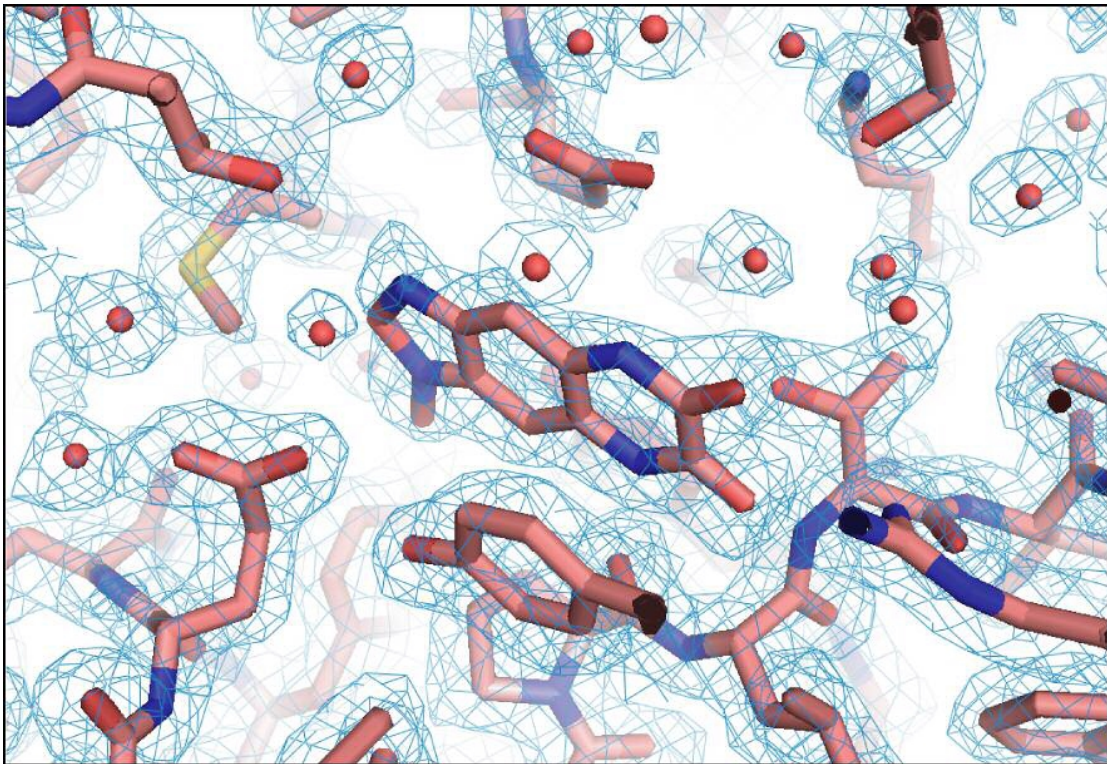


Figure S2. Fo-Fc omit electron density for FQX and surrounding ligands contoured at 1.0 sigma. At 1.5 sigma, the electron density shows the good shapes for the side chains and aromatic rings and has reliable solvent peaks. At 3.0 sigma contouring, there is significant loss of information.

Table 1. Statistics for X-ray Data Collection and Refinement.

	S1S2J-FQX
Data Collection Statistics	
Space group	$P2_12_12_1$
Cell constants a/b/c (Å)	62.34
	92.26
	195.24
# Molecules/ASU	4
Total Reflections	191553
Unique Reflections	90300
Reflections used refinement	90300
R-merge(%)[a,b]	5.6 (51.2)
Redundancy	2.1
I/ σ (I)[b]	8.0 (2.0)
Completeness (%)[b]	97.0 (92.8)
Refinement Statistics	
Resolution (Å)	1.87
R-factor (%)[c]	21.2
R-free (%)[d]	22.0
B-factors	
Protein	17.677
Ligand	23.172
Solvent	24.521
R.M.S.D bonds	0.017
R.M.S.D angles	1.7
Waters	875
Matthews Coefficient (Da-1)	2.34
Solvent content (%)	47.4
Ramachandran plot (%)	
Most Favored	94%
Allowed	6%
PDB code	3BKI

ASU=asymmetric unit

[a] R merge (%)= $\sum_{hkl} |<I>-I| / \sum_{hkl} |I|$

[b] Values in parentheses refer to the highest resolution shell (1.87-1.97 Å).

[c] R factor (%)= $\sum_{hkl} ||Fo| - |Fc| / \sum_{hkl} |Fo|$

[d] 5% of the reflections were set aside for the calculation of the R free value.

References

(1) Chen, G. Q.; Sun, Y.; Jin, R.; Gouaux, E. Probing the ligand binding domain of the GluR2 receptor by proteolysis and deletion mutagenesis defines domain boundaries and yields a crystallizable construct. *Protein Sci* **1998***, 7*, 2623-2630.

(2) Holton, J., Alber, T. Automated protein crystal structure determination using ELVES. *Proc Natl Acad Sci U S A.* **2004***, 101*, 1537-1542.

(3) Armstrong, N.; Gouaux, E. Mechanisms for activation and antagonism of an AMPA-sensitive glutamate receptor: crystal structures of the GluR2 ligand binding core. *Neuron* **2000***, 28*, 165-181.

(4) Brunger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P. et. al. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta. Crystallogr. D. Biol. Crystallogr.* **1998***, 54*, 905-921.

(5) Emsley P.; Cowtan K.; Coot: model-building tools for molecular graphics *Acta Crystallographica Section D-Biological Crystallography* **2004**, *60*, 2126-2132

(6) Engh, R. A.; Huber, R. Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Cyrstallographica* **1991***, A47*, 392-400.

Chapter 2

Re-engineering Androgen Receptor Towards Transcriptional Activation with Novel Steroids

Introduction

Our goal was to design a steroidal ligand-hormone receptor pair such that the novel ligand does not bind to endogenous androgen receptors (AR), and the designer receptor does not bind to endogenous androgens. The application for this orthogonal ligand-receptor pair is to study male specific behavior in mouse. Gender-specific behaviors in vertebrates such as mating and aggression are regulated by gonadal hormones. Thus, testosterone is required for male behaviors whereas estrogen and progesterone are essential for female behaviors. Steroid-activated zinc finger transcription factor receptors are expressed in the brain, suggesting that gonadal hormones directly influence neurons. Heterologous agonist-receptor pairs designed after high resolution structural studies have yielded spectacular insights into signaling by kinases and G-protein coupled receptors. We want to extend this idea to the realm of steroid receptors and behavior. Androgens acting through the androgen receptor not only control development and gene expression but also gender-specific behaviors such as mating and aggression. Little is known about neural responses evoked by steroids such as testosterone because neural studies are difficult due to the pleiotropic effects of these hormones throughout the body. The orthogonal ligand-receptor pairs generated from this project should permit inducible and tissue-specific manipulation of AR signaling in the future. These and later experiments will lead to mechanistic

insights into the role of steroid regulation of neural circuits that mediate gender-specific behaviors.

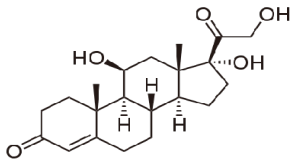
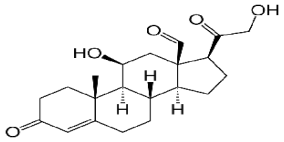
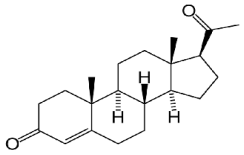
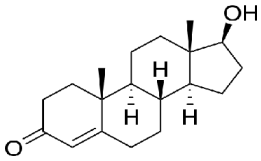
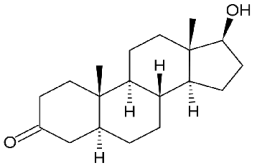
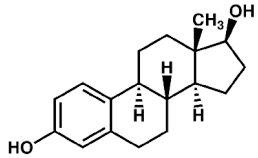
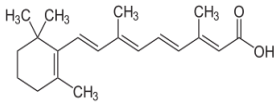
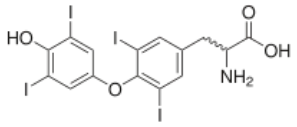
Three approaches were used before acquiring androgen receptor mutants that transcribed with a steroid previously unable to activate transcription. The first attempt was rational design. The C-19 position of testosterone is positioned close to two methonines. Rationally, if the methionine is mutated to an aniline this will accommodate an additional propyl group at the C-19 position. Custom synthesis of 19PT, an analog of testosterone with an additional propyl group at the C-19 position, was performed by Blake Peterson, Penn State. There was no noticeable transcriptional activation with the two mutants M742A and M745A in cell based assays. The second attempt utilized the ORBIT program (Stephen Mayo – CalTech). 19PT was modeled in to the binding pocket of AR LBD and the side chains were systematically repacked to accommodate 19PT. The resulting suggested AR LBD mutants were created via site-directed mutagenesis and tested in a GAL-4_AR-LBD transcription assay. However, the best mutant only had 5% transcription. Finally, a directed evolution yeast-based selection assay was used to select from ~10^9 AR mutants to produce 9 AR mutants that activate transcription with 100 nM gestrinone, an endometriosis drug, as a proof-of-concept. Characterization and follow-up studies are currently under investigation.

Background

Steroid Nuclear Receptor Family

Since the cloning of the first nuclear receptor, human glucocorticoid receptor (GR) (Hollenberg et. al., 1985), over 20 years ago by Evans and coworkers, there has been a huge accumulation of data on NR-dependent transcriptional regulation (Mangelsdorf et. al., 1995; Aranda & Pascual, 2001; McKenna & O'Malley, 2002a, b; Nagy & Schwabe, 2004). Steroid hormones are among the most fundamental signaling molecules in nature and are responsible for the regulation of development, reproduction, metabolism, and response to environmental cues (Mangelsdorf et. al. 1995). Steroid hormones bind to a small family of cellular and nuclear receptors, the nuclear hormone receptors (NRs). NRs are ligand-dependent and -independent transcription factors that are highly conserved evolutionarily from invertebrates to higher organisms (Keay and Thornton, 2009). The NR superfamily regulates a complex network of genes that coordinate nearly all the activities of homeostasis, growth, and reproduction (Novac & Heinzel, 2004; Margolis et. al., 2005). The human genome includes 48 NR genes. Alternative splicing and promoter usage of these 48 genes give rise to 75 currently known NR proteins (Lander et. al., 2001; Venter et. al., 2001; Robinson-Rechavi et. al., 2001, 2003a; Escriva et. al., 2004). NRs are ligand (hormone) inducible transcription factors that, with the assistance of

auxiliary proteins (coregulators), regulate the expression of their target genes in a temporal and tissue-specific manner (Aranda & Pascual, 2001; McKenna & O'Malley, 2002a, b; Novac & Heinzel, 2004). The conservation between NRs suggests that they all come from a common ancestor by gene duplication and divergence (Escriva et. al., 2004; Thornton & Kelly 1998). NRs are characterized by their modular structure. This consists of a highly variable amino-terminal domain (NTD). The NTD region can range in size from being 6% of the total protein, as for the vitamin D receptor (VDR), to over 50% of the total protein, as for AR. NRs also have a highly conserved central DNA-binding domain (DBD) and conserved ligand-binding domain (LBD). Within the NR superfamily, the steroid receptors make up a subfamily that responds to hormones based in part on steroid scaffolds (Table 1); these include the estrogen receptor (ER), progesterone receptor (PR), androgen receptor (AR), glucocorticoid receptor (GR), thyroid receptor (TR), vitamin D receptor (VDR), and mineralocorticoid receptor (MR). The endogenous hormones that activate the steroid nuclear receptor family are very similar in structure, which is in stark contrast to the enormous differences in physiological effects (Table 1). All steroids are lipophilic molecules derived from cholesterol. (Bolander, 2004; Alberts et. al., 2002).

Nuclear Receptor	**Gene**	**Primary Endogenous Hormone**	**Structure**
Glucocorticoid	NR3C1	Cortisol	
Mineralocorticoid	NR3C2	Aldosterone	
Progesterones	NR3C3	Progesterone	
Androgen	NR3C4	Testosterone	
		Dihydrotestosterone	
Estrogen	NR3A1 (α) NR3A2 (β)	Estradiol	
Retinoic Acid	NR1B1	All-trans-retinoic acid	
Thyroid	NR1A2	Thyroxine (T4)	

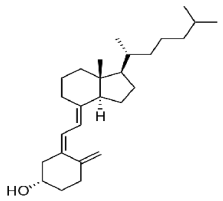
Vitamin D	NR1I1	Cholecalciferol	

Table 1: Steroid Nuclear Receptors and Endogenous Hormones

Structure Of The Androgen Receptor

Genomic Structure and Organization

The genomic structure/organization of the *AR* gene is conserved in the mammalian kingdom from mouse to man (Gelmann, 2002) (Figure 1). Human AR is encoded by a single 180 kbp gene found on the long arm of chromosome X at Xq11-12 (Lubahn et. al., 1988; Brown et. al., 1989). The mRNA transcript is 10.6 kbp long and has an open reading frame of 2757 bp, which codes for the eight exons of *AR*. Between 1988 and 1989 several groups cloned the human *AR* complementary DNA (cDNA) (Chang et. al., 1988; Lubahn et al, 1988; Trapmann et. al., 1988; Tilley et. al., 1989).

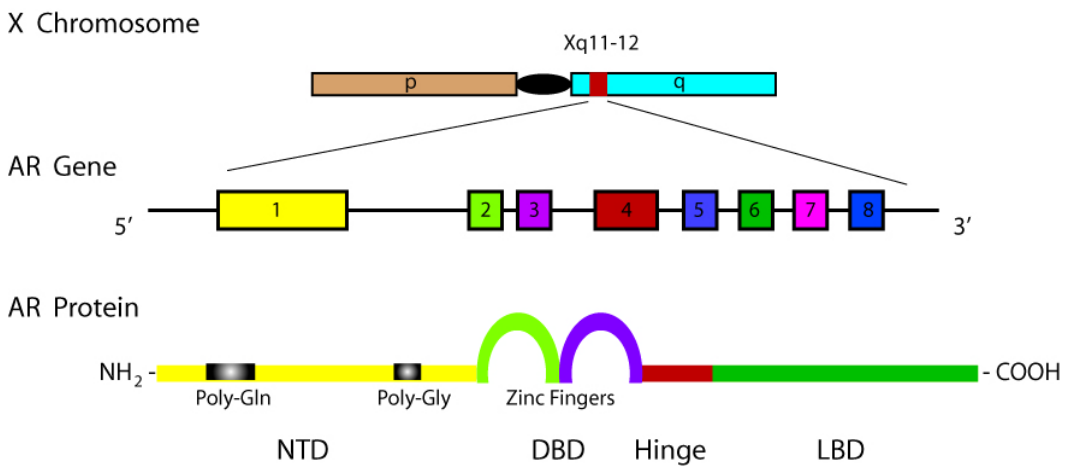


Figure 1: Structural organization of the human AR gene. The AR gene is located at Xq11-12. The exons (1-8) are colored and the relationship to the functional domains they encode are shown. NTD, amino-terminal domain; DBD, DNA-binding domain; LBD, ligand-binding domain.

Structural Features of Androgen Receptor

The AR protein of human, rat and mouse are all approximately 99 kDa (un-phosphorylated) or 110-kDa (post-transcriptionally phosphorylated). Human AR protein consists of 919 amino acids and is organized in three main structural domains (Figure 1): N-terminal domain (NTD), DNA binding domain (DBD), ligand-binding domain (LBD). The DBD and LBD are 100% conserved, while the hinge and NTD are about 70-80% conserved. The N-terminal domain (NTD) is the least homologous in sequence and most variable in size domain among the members of the steroid receptor family. This domain is mainly involved in the regulation of target gene transcription as well as in transcriptional regulation via protein-protein interactions

with other transcription factors (Rundlett et. al., 1990; Jenster et. al., 1991; Simental et. al., 1991; Palvimo et. al., 1993). The DNA binding domain (DBD) is the most highly conserved part of the receptor molecule, which determines the specificity of AR interaction with DNA (Freedman et. al., 1992; Berg et. al., 1989). DBD consists of two zinc clusters: one is involved in direct DNA-binding and contains the P-box for the specific recognition of the androgen-response element, while the other is implicated in protein-protein interactions and serves as a stabilization unit for the dimerization of the two receptor molecules. Between the DBD and the LBD lies the hinge region, which contains the major part of the AR nuclear targeting signal and mediates the transfer of AR from the cytoplasm to its site of action in the nucleus (Jenster et. al., 1991; Simental, et. al., 1991; Zhou et. al., 1994). The ligand-binding domain (LBD) functions principally by specific, high-affinity binding of androgens. In addition, the LBD is also involved in nuclear localization, receptor dimerization and interaction with other proteins (Kuiper et. al., 1989). The amino acid sequence of this region displays about 50% homology with the corresponding residues in glucocorticoid, mineralcorticoid and progesterone receptors (Quigley et. al., 1995). Activation function 2 (AF-2) is responsible for agonist induced activity (activity in the presence of bound agonist). AF-2 binds either the N-terminal FXXFL motif intra-molecularly or co-activator proteins (containing the LXXLL or preferably FXXFL motifs) (Dubbink et.al., 2004).

Many crystal structures of the human AR ligand-binding domain (hARLBD) have been solved in complex with DHT (Sack et. al. 2001) and with chemically modified steroids such as the agonist metribolone (R1881) (Matias et. al. 2000), the corticosteroid agonist 9α-fluorocortisol (Matias et. al. 2002), and the designer androgen, tetrahydrogestrinone (Pereira de Jésus-Tran et. al., 2006). In addition, structures of the liganded hARLDB have been determined in complex with a peptide derived from physiological co-activators (He et. al. 2004; Hur et. al. 2004; Estebanez-Perpina et. al. 2005). In all these complexes, the LBD adopts the same fold, mainly composed of α-helices arranged as a three-layered anti-parallel α-helical sandwich, a fold common to all the NRs (Wurtz et. al. 1996). These structures show that the ligand binding pocket (LBP) is mainly composed of hydrophobic residues, the side chains of which can easily adopt variable positions in order to better fit the hydrophobic core of the steroid and stabilize it. The capacity of the androgen receptor to modify the shape of its ligand binding pocket in order to accommodate ligands with different structures is well demonstrated by the variation of the LBP volumes in different complexes. The size of the LBP was measured for each complex (Testo = 584 $Å^3$, DHT = 582 $Å^3$, THG = 605 $Å^3$) and is proportional to the volume of the ligand itself (Testo = 252 $Å^3$, DHT = 249 $Å^3$, THG = 275 $Å^3$), each ligand occupying <50% of the available space (Pereira de Jésus-Tran et. al., 2006). The rest of the space is occupied by the hydrogen atoms belonging to the amino acids forming the cavity and to the ligand. A similar observation was seen with the human enzyme 17β-hydroxysteroid de-hydrogenase type I (HSD10) , known to be specific to

estrogens but which is able to bind other steroids by rearranging some side chains of its steroid-binding site (Blanchet et. al., 2005).

The Androgen Receptor and Androgens

Endogenous androgens, testosterone (T) and dihydrotestosterone (DHT) exert their effects by mediating the differentiation and development of the normal male phenotype via the androgen receptor (AR) (McPhaul et. al., 1999). Androgen receptor relays the signals of androgens to the basal transcription machinery in the nucleus (Quigley et. al., 1995; Gelman, 2002; Lee & Chang, 2003). Like all NRs, AR has a conserved molecular structure, with each domain playing an important role in AR function and signaling. This is either via intra-receptor interactions or via functional interactions with androgen response elements and/or co-regulatory proteins (Heinlein & Chang, 2002; Glass & Rosenfeld, 2000; McKenna et al, 1999a, b; McKenna & O'Malley, 2002a, b). In the absence of ligand, the AR resides in the cytoplasm (Georget et. al., 1997). Hormone binding induces a conformational change of the receptor and allows its dimerization and trans-location into the nucleus, where it initiates transcription through specific interactions with the transcription machinery (Meehan et. al., 2003). Disturbances in AR functionality caused by receptor mutation, disrupted DNA interactions, or altered co-regulator interactions appear to be linked to a range of diseases including androgen insensitivity syndrome (AIS) and prostate cancer (McPhaul, 1999, 2002; Arnold & Isaacs, 2002; Abate-Shen & Shen, 2000; Parkin et. al., 2005).

Androgen receptor is expressed in fetal tissues as early as 8 weeks of gestation, before the onset of androgen action, and is activated in a ligand-dependent manner to coordinate expression of suitably responsive genes. Testosterone serves as a substrate for two metabolic pathways that produce sex steroids. Testosterone is reduced by 5α-reductase to produce DHT or be aromatized to generate estrogens. Dihydrotestosterone (DHT) is essential for the development of the penis, scrotum and prostate (Nef and Parada, 2000). In the absence of testosterone production or in the presence of estrogens, these male determining structures regress and female sexual organs form. In the absence of androgens or faulty AR function genetically male embryos develop a female phenotype. Therefore, the synthesis of each of these steroids in developing male and female embryos must be subjected to a regulation that maintains the delicate balance between androgens and estrogens (Nef and Parada, 2000). Testosterone and DHT control the development, differentiation and function of the male reproductive and sex tissues, such as seminal vesicles, epididymides and prostate. Other organs influenced by androgens include skin, skeletal muscle, bone marrow, hair follicles and behavioral centers of the brain (Quigley et. al., 1995; Gelman, 2002).

Testosterone is the main androgen in men and the testes produce about 80-95% of circulating testosterone, while the adrenal glands produce the remaining 5-20% (Shen & Coetzee, 2005). In human target tissues the concentration of testosterone can range from 100 nM up to 1 μM as found in the intratesticular fluids,

but the percentage that is active remains unknown (Jarow et. al., 2005). In women the major source of androgens is not from the adrenal glands, but is from ovary derived estrogens converted to testosterone (Shen & Coetzee, 2005).

Anabolic Steroids

Anabolic steroids are drugs that mimic the effect of endogenous steroids, testosterone and dihydrotestosterone. In the 1930s, anabolic steroids were first isolated, identified, and synthesized. Today, these drugs are used therapeutically to stimulate bone growth and appetite, induce male puberty, and treat chronic wasting conditions, such as cancer and AIDS. Use of anabolic steroids in sports and bodybuilding has been banned by all major sporting bodies. A novel chemically modified steroid, tetrahydrogestrinone (THG), has appeared as a doping agent. A potent androgen and progestin (Death et. al., 2004), THG is produced by the hydrogenation of gestrinone, a progestin used to treat endometriosis (Dawood et. al. 1997), and has been identified as the first true "designer steroid," being custom produced to evade detection (Catlin et. al., 2004). It was undetectable in urine by standard anti-doping tests until a specific test was developed (Catlin et. al., 2004). Using a genomic assay, THG has been shown to modulate hundreds of genes in a time-dependent fashion almost identical to DHT (Labrie et. al. 2005).

Protein Engineering

Nearly 30 years ago, site-directed mutagenesis was first used to modify enzymes of known structure and mechanism. Since then, this method has contribute

insights into catalysis, specifically, stability and folding of proteins (Brannigan et. al., 2002; Jackel et. al., 2008). There are two general strategies for protein engineering, rational design and directed evolution. In rational protein design, detailed knowledge of the structure and function of the protein is used to make desired changes. This generally has the advantage of being inexpensive and technically easy, since site-directed mutagenesis techniques are well-developed. However, it can be extremely difficult to predict the effects of various mutations. In directed evolution, random mutagenesis is applied to a protein, and a selection regime is used to pick out variants that have the desired qualities. The great advantage of directed evolution is that it requires no prior structural knowledge of a protein, nor is it necessary to be able to predict what effect a given mutation will have. The results of directed evolution experiments are often surprising in that desired changes are often caused by mutations that were not expected to have that effect (Jackel et. al., 2008). The drawback is that they require high-throughput assays, which is not feasible for all proteins. Large amounts of recombinant DNA must be mutated and the products screened for desired qualities. The sheer number of variants often requires expensive robotic equipment to automate the process (Looger et. al., 2003).

Chemical Genetics & Orthogonal Ligand-Receptor Pairs

Chemical genetics is an approach that uses small molecules to alter protein function – directly, in real time, rather than indirectly by manipulating their genes. It is used to identify which proteins regulate different biological processes, to

understand in molecular detail how proteins perform their biological functions, and to identify small molecules that may be of medical and/or commercial value.

The approach of chemical genetics has been widely used to control various protein functions and cellular processes (Buskirk et. al., 2005). For example, this approach has been used in studying GTP regulatory proteins to create nucleotide specificity (Hwang et. al., 1987). It has been suggested that the chemical genetic approach is superior to the traditional genetic approach, such as gene knockout, in studying protein functions, (Strausberg et. al., 2003) and has advantage for developing drugs against genetic diseases (Koh et. al., 2005). Highly-specific ligands for thyroid hormone nuclear receptors may have potential to restore mutation-caused genetic diseases (Hashimoto et. al., 2005). Functional orthogonal ligand−receptor pairs have been used for regulation of estrogen receptors (Shi et. al., 2002; Tedesco et. al., 2001), Src family protein kinases (Liu et. al., 1998), and protein methyltransferases (Lin et. al., 2001).

Re-engineering the Estrogen receptor ligand-binding domain.

Koh and colleagues explored a protein engineering strategy of manipulating polar/charged residues across the ligand receptor interface of estradiol (E2) and the estrogen receptor (ER) (Koh et. al., 2005; Shi et. al., 2002). Carboxylate-functionalized E2 analogues can activate ERα(Glu353→Ala) and ERβ(Glu305→Ala) with very large selectivities, demonstrating that their design strategy could be

extendable to other members of the steroid hormone receptor family. Neutral E2 analogues were found to complement ERα(E353A) with similar potencies but with generally lower selectivities. This suggested that the high selectivity observed with ligand−receptor pairs generated by exchanging charged residues across ligand−receptor interfaces is only due in part to their complementary shapes and that appropriate introduction of charged functionality on the ligand can provide substantial enhancement of selectivity by decreasing the engineered ligands affinity for the endogenous receptor.

Katzenellenbogen and colleagues created an hERα ligand-receptor pair with a directed evolution approach that had enhanced androgen specificity and affinity (Tedesco et. al., 2001; Chen et. al., 2004; Chockalingam et. al., 2005). They developed a sensitive yeast two-hybrid system to screen for hERα mutants with increased trans-activation potency toward testosterone. After two rounds of directed evolution, they identified five hERα mutants with dramatically improved trans-activation potency toward testosterone in both yeast and mammalian cells. These variants showed up to 7,600-fold improvement in the binding affinity for testosterone and only slightly reduced affinity toward 17beta-estradiol. Detailed analysis of these evolved variants and a few site-directed mutants generated de-novo led to several unexpected findings: Only two beneficial mutations were needed to create hERα mutants with near nanomolar affinity for testosterone. Some beneficial mutations were synergistic, context-dependent, or non-additive. Of the five identified beneficial

mutations, four of them were not in the ER ligand binding pocket and yet exerted important action on ligand specificity. The single ligand-contacting mutation E353Q plays a dominant role in discriminating androgens and estrogens.

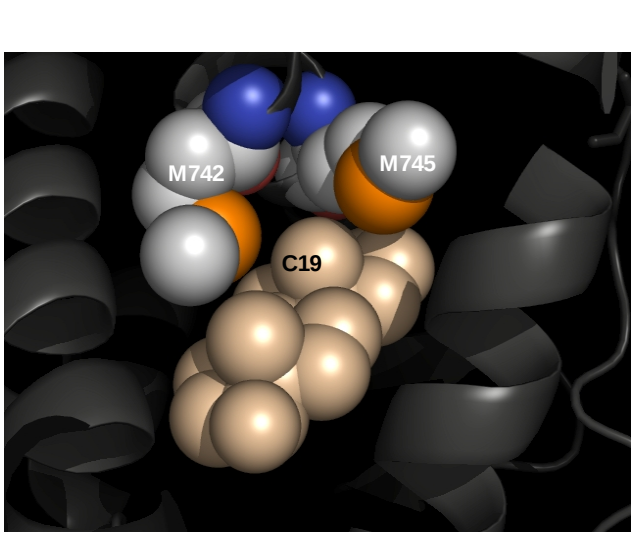
Results

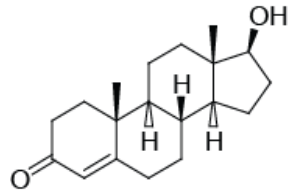
Rational Design Approach Of AR LBD With 19PT

Design and Synthesis of 19PT

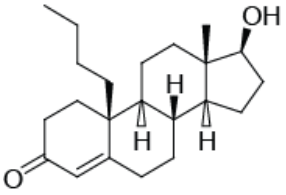
We chose to design an androgen that is very close to the natural hormone so that it would be specific when delivered in vivo. The properties of the novel androgen would be to bind solely to an androgen receptor, mutant AR that was engineered over the natural receptor, AR, by at least one hundred fold. The novel androgen would not bind to other steroid receptors as we would use the endogenous hormone for the receptor, DHT, as the scaffold. We also are constrained by the difficulty of doing synthetic chemistry on the steroid scaffold. Using high resolution details of our AR-LBD/DHT crystal structures in an exacting series of atom exchanges, we identified the C19 atom, the single position on the backbone of the natural ligand DHT, as the site of chemical substitution (Figure 2). We designed a compound 19-propyltestosterone (19PT; $C_{22}H_{34}O_2$, M.W. = 330.5) carrying a propyl extension at C19. This compound compares with testosterone in having a critical methyl group, C19 at C10, substituted by a butyl. Because steroid hormones are

folding keys for their receptors, this extra mass would insure tight binding only to an accommodating engineered steroid receptor. Custom synthesis of 19PT was carried out early in 2006 by Dr. Blake Peterson (Department of Chemistry, Penn State). This compound represents a novel androgen that is hypothesized to be selective activator of mutant AR versus wild type AR. Figure 2 shows two methionine side chains, M742 and M745 to be adjacent the C19 position. We hypothesized that 19PT will be the exclusive agonist partner of AR_M742A, or AR_M745A.





Testosterone



19PT

Figure 2: A: Space-filling model of testosterone (salmon), M742, M745 (light gray, orange). B: Chemical structure of testosterone C: Chemical structure of 19PT.

Initial Binding studies of 19PT

AR without hormone is unstable. We conducted the comparative protein expression studies in *E. coli*, where the recombinant AR ligand-binding domain

(LBD; a.a. 662-919) was expressed in the presence of either DHT or 19PT. As predicted from the engineered models, AR expression with 19PT was negligible as compared to AR expression with DHT. Further attempts to isolate AR LBD with 19PT failed (Figure 2). The *E. coli* expression test is indicative of the very low binding affinity of 19PT to AR. Similarly AR can not be expressed with any natural steroid ligand other than DHT or testosterone. 19PT is shown below to be inactive with wild type AR in transcription activation assays in HeLa cells (data not shown).

We hypothesize that the propyl extension of 19PT interferes with 2 methionine side chains of the protein backbone, M742 and M745, which prohibits binding of the compound and the subsequent folding of the receptor in *E. coli*.

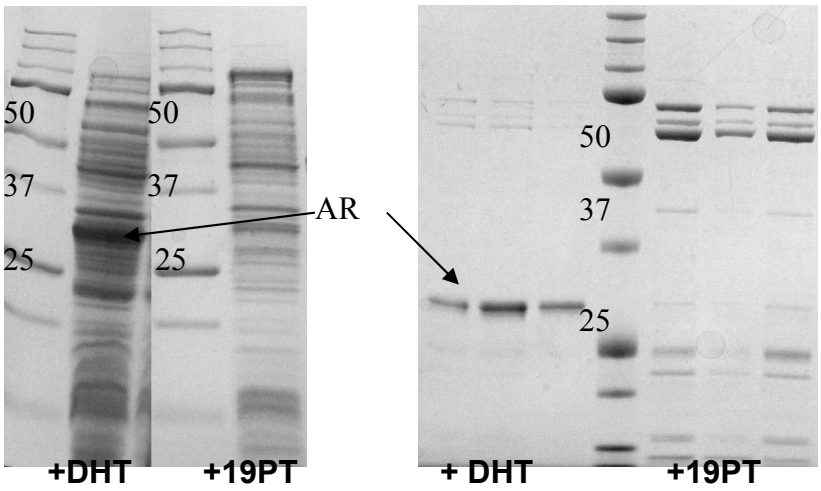


Figure 3: SDS-PAGE of AR LBD expression and purification in the presence of DHT versus 19PT. Lanes 1,2. Cell lysate with recombinant AR LBD expression, DHT (1) or 19PT (2) was added during expression. AR/DHT is well expressed, while AR/19PT is not. MW of AR LBD is 33 kDa. Lanes 3-5. AR/DHT protein fractions

purified using Talon matrix. Lanes 6-8. No AR protein present in the Talon matrix purification step, where 19PT was used.

Androgen Receptor Mutants Designed By ORBIT

We collaborated with Stephen Mayo and Heidi Privett at CalTech to create a second round of AR mutants that could bind/activate transcription with 19PT. ORBIT (Optimization of Rotamers By Iterative Techniques) is a computational design algorithm based on physical chemical potential functions and stereochemical constraints designed by the Mayo lab (Dahiyat and Mayo,1997a, b; Dahiyat et. al., 1997; Street and Mayo, 1998; Pierce et. al., 2000).

Five AR mutants were identified with ORBIT. Four AR mutants contain three point mutations and one mutant contains one point mutation. All the residues contact the ligand, 19PT, except V715, which is in the second shell of residues surrounding the hormone-binding pocket. These mutants were designed by creating a library of small molecule rotomers of 19PT, modeling in library of rotomers into binding pocket (Figure 4b), "repacking" binding pocket with residues that accommodate small molecule, and, finally, ranking mutants based on energy of Van der Waals packing interactions.

AR_V715L_W741F_M745A
AR_V715I_W741F_M745A
AR_V715I_W741H_M745A
AR_V715L_W741L_M745A
AR_W741A

Table 2: AR LBD Mutants Designed by ORBIT

Point mutations pertaining to the ORBIT mutants were introduced in a GAL4-AR_LBD (662-919) vector using site-directed mutagenesis. The single, double, and triple mutants were isolated and sequenced to ensure placement of proper mutations. The mutants were analyzed with a Gal4 transcription assay (Figure 5). The 25 mutants had very low levels of transcription with 10 μM 19PT. At most, AR_V715I had 5% transcriptional activation. With 10 nM DHT, the 25 mutants had variable levels of transcription compared with wild-type AR_LBD. Further analysis of a select group of mutants revealed that there was variation in protein expression that lead to the different levels of transcription with DHT (data not shown).

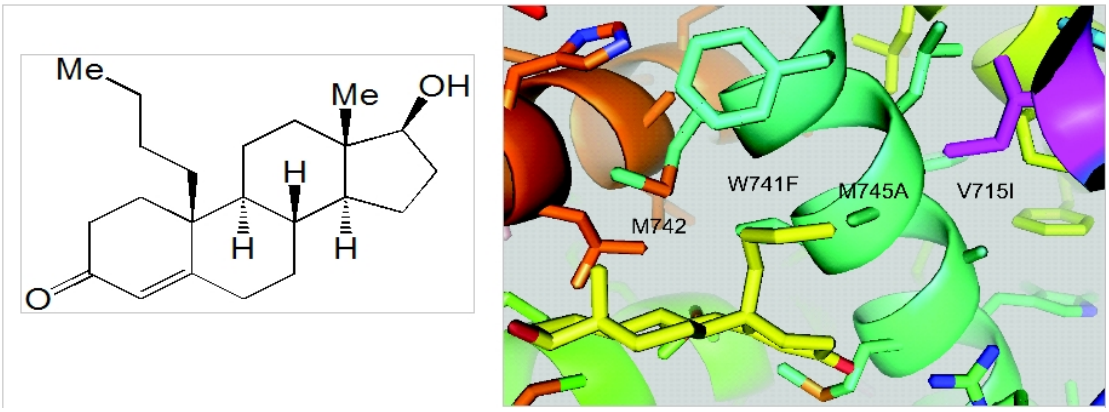


Figure 4: A. The novel androgen, 19PT, with the propyl extension at C-19.

B. Computational representation of a triple mutant (V715I_W741F_M745A) of the androgen receptor ligand binding pocket (ribbons) with 19PT bound (yellow). These mutations show that the butyl group of 19PT replaces the methionine. While V715I and W741F compensate for the changes in increased and reduced space, respectively.

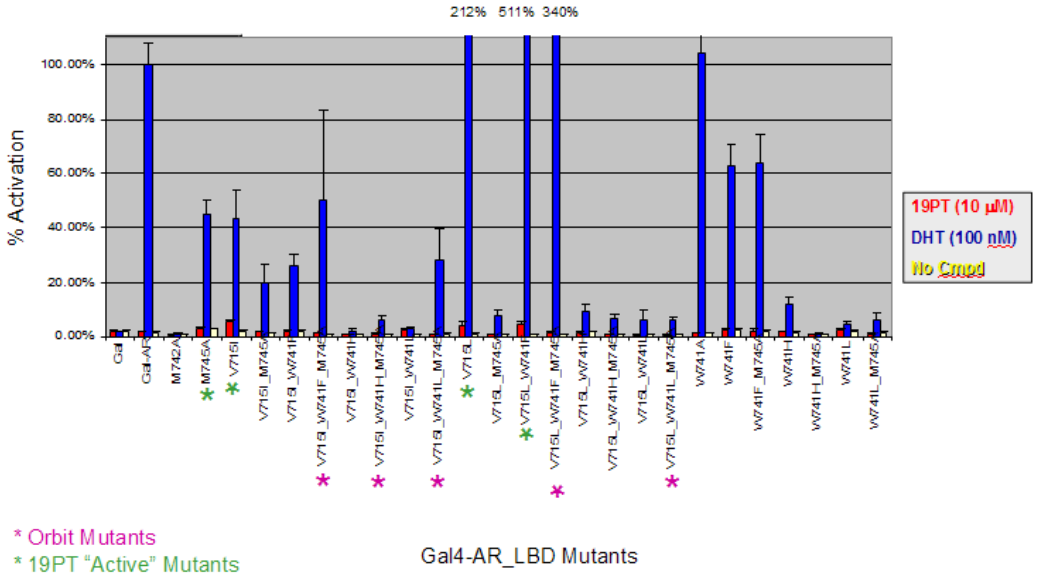
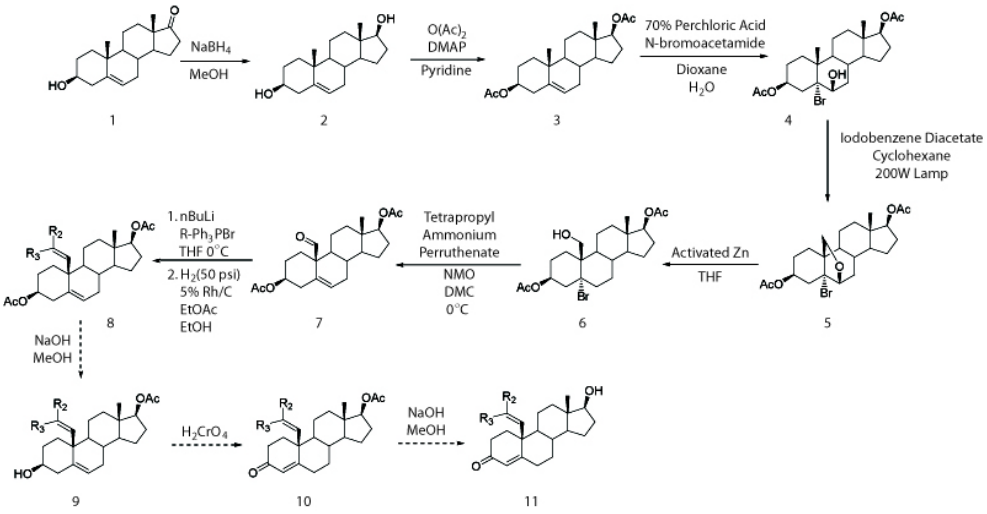


Figure 5: Transcriptional Data of ORBIT AR LBD mutants

In-house Partial Synthesis Of 19PT

This proposed synthetic scheme for 19PT is based on published literature (Scheme 1) (Blizzard et. al., 2006a; Blizzard et. al., 2006b). The C-3 and C-17 of commercially available androstenediol is first protected as acetates. Functionalization of the 5,6-olefin is accomplished by treating the protected diol intermediate with N-bromoacetamide, in the presence of perchloric acid. The product of this reaction has an axial hydroxyl group at C-6 of the steroid nucleus which serves as the handle for oxidation of the C-19 methyl group. This can be accomplished by photolyzing a mixture of the alcohol, iodobenzene diacetate, and iodine in cyclohexane. Reduction of the resulting cyclic ether with activated zinc dust regenerates the 5,6-double bond and affords a 19-hydroxy steroid. (To obtain the alkanyl analogs, an additional step is required at this stage to reduce the olefin to an alkane. The remaining synthetic steps can be preformed on the olefin and alkane analogs.) The 19-hydroxy steroid is the oxidized to an aldehyde, using tetrapropyl ammonium peruthenate (TPAP) and N-methyl morpholine N-oxide (NMO) in dichloromethane. Under Wittig conditions, the 5,6-olefins are obtained from the aldehyde. To achieve the testosterone-like steroids, 5,6-olefins are first treated with sodium hydroxide and methanol to selectively de-acylate the C-3 acetate, followed by oxidation of the resulting C-3 hydroxyl to a ketone. The final step is de-protection of the C-17 acetate.



Scheme 1: Detailed Synthetic Route of 19PT and Analogs

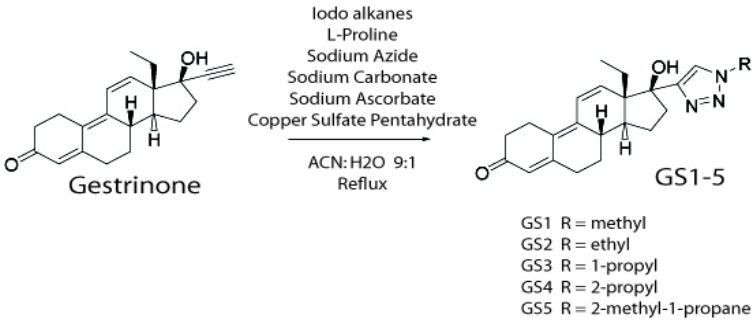
The synthesis was carried out until the Wittig Reaction. Sufficient quantities of compounds **1-6** with high yields (80-99%) was accomplished. The yield for compound **7** ranged from 20-30%. Many attempts were made with the Wittig reaction however, the synthesis of 19PT was stopped due to changes in priority for this orthogonal ligand-receptor pair project.

Gestrinone And Triazole-analogs

Since the synthesis of 19PT was long and complicated, a search was made for for a commercially available steroid that had little or no activity with androgen receptor and could be modified to obtain additional analogs. Gestrinone, an endometriosis drug with very little activity with androgen receptor (Dawood et. al., 1997) and a terminal alkyne at the C-17 position, met this criteria . The main targets

of gestrinone are progesterone and estrogen receptor (Dawood et. al., 1997). Gestrinone is also the precursor to tetrahydrogestrinone, the most potent anabolic steroid (Death et. al., 2004; Catlin et. al., 2004). I hypothesized that bulky groups added to the alkyne at the C-17 position of gestrinone would yield steroids that were unable to activated transcription with AR.

Synthesis of triazole analogs of gestrinone was completed using *in situ* click chemistry in one step, utilizing the alkyne at the C-17 position (Scheme 2) (Feldman et. al., 2004). The reaction is very difficult because low molecular weight alkyl azides are highly unstable and must be synthesized *in situ* from commercially available iodoalkanes. GS1-5 were synthesized with varying alkyl groups (Scheme 2). Other triazole analogs with larger alkyl groups were proposed but computational docking of these analogs showed that the larger alkyl groups could not be accommodated in the ligand-binding pocket of AR (Kris Kuchenbecker, data not shown). Using the Gal4-AR LBD transcription system in HeLa cells, the hypothesis was confirmed (Figure 6). There is no transcriptional activation with GS1-5 and only negligible transcription with gestrinone.



Scheme 2: Synthesis of Triazole Analogs of Gestrinone Via Click Chemistry

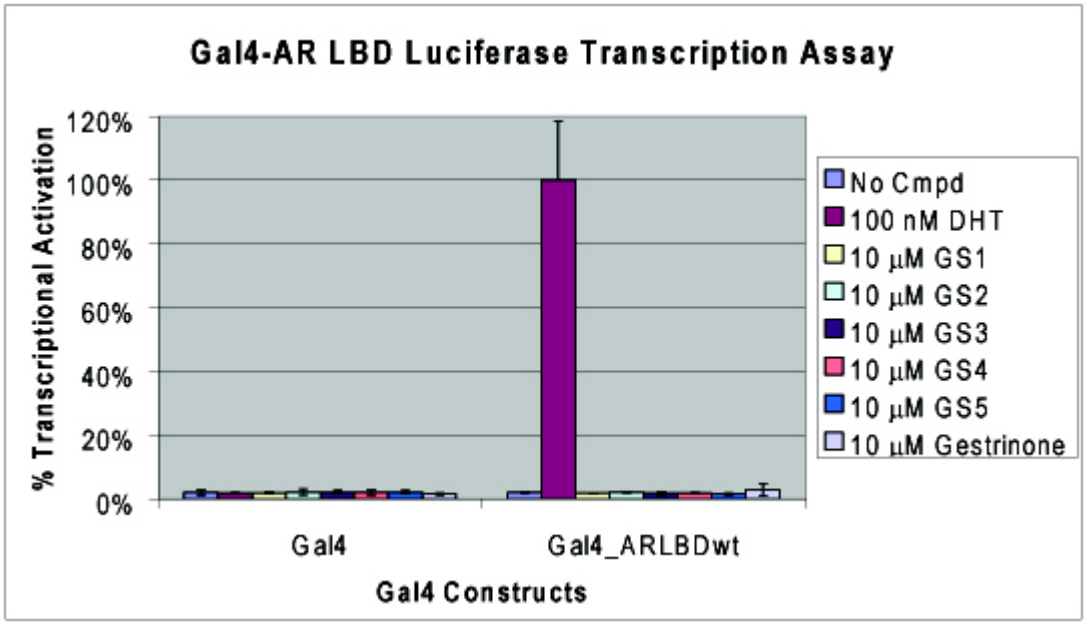


Figure 6: Transcriptional activation was measured using Gal4-AR LBD transcription system in mammalian cells (HeLa). 10 mM gestrinone and 10 mM GS1-5 derivatives do not activate transcription of Gal4-AR LBD, which is more sensitive than full length AR. There is no transcriptional activation with GS1-5 and negligible transcription with gestrinone.

Directed Evolution Of Androgen Receptor

After synthesizing GS1-5, we had to decide on which method to use to create an AR magic pair. I did not want to go back to rational design or ORBIT. We found a useful yeast assay developed by Marc Cox, UTEP. He was using this yeast assay to study FKBP52 and FKBP51 function with steroid nuclear receptors. Altering the assay allowed me to use it as a directed evolution assay with full-length androgen receptor. We are using this genetic selection assay in *S. cerevisiae* to reveal gain of function activity from mutated AR. This yeast-based selection assay is very powerful

because only viable AR mutants that are able to transcribe the *HIS3* gene will allow the yeast to propagate and form colonies.

Step one consists of error prone PCR generation of mutant AR plasmids. Random mutations were introduced between residues 566-919 of the full length AR (Figure 7). Approximately 10^9 AR mutants were generated using error prone PCR. The PCR insert and linear vector (p464GPD-AR) are co-transformed into an engineered strain of *S. cerevisiae*, which recombines the DNA segments via homologous recombination.

Step two is the directed evolution yeast-based selection assay. Integrated into the yeast genome are six hormone response elements (HRE), promoters (PCYC1), *HIS3* and *lacZ* genes (Figure 7). FKBP52, a nuclear receptor co-chaperone, increases the sensitivity of the assay (Pratt et. al. 1997, 2003, 2004). Under the constraints of the selective medium, if the AR mutant is viable (stable, properly folded) and if the steroid is able to bind and activate transcription, the *HIS3* gene will be transcribed allowing the yeast to have all of the necessary nutrients to grow.

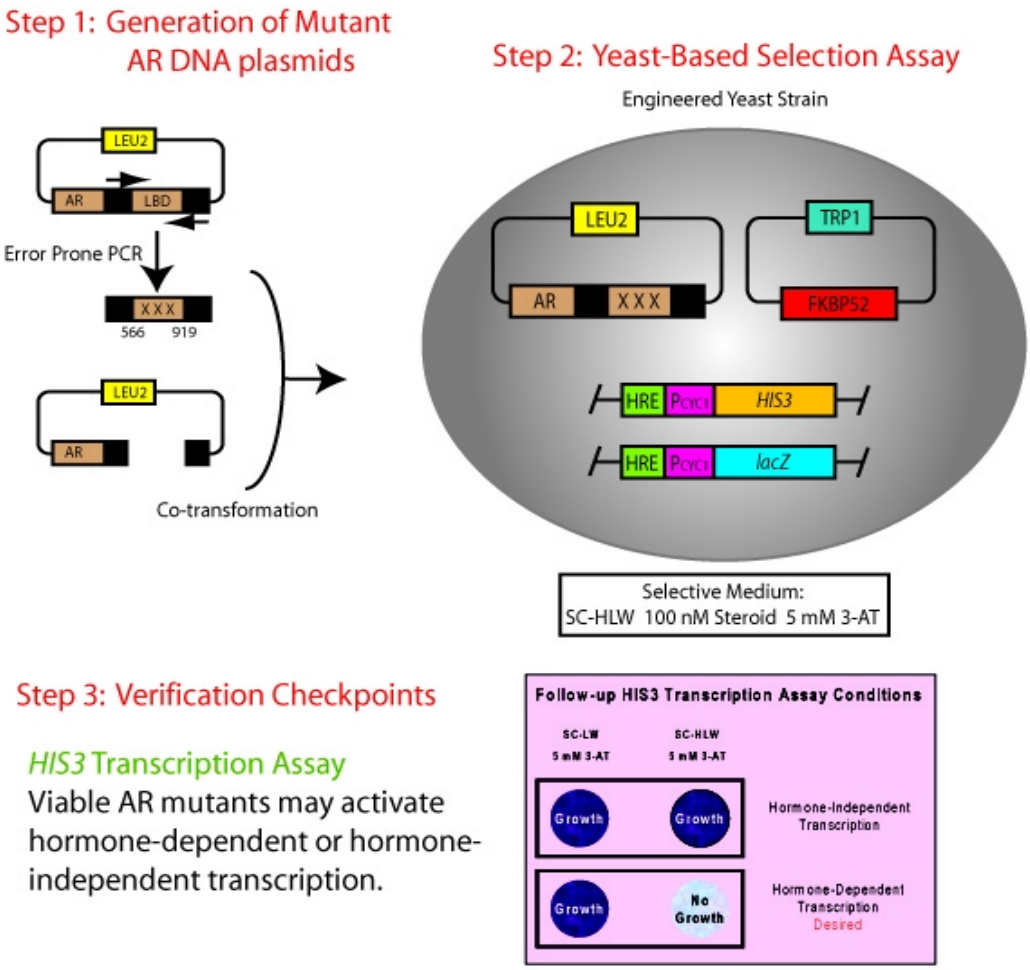


Figure 7: Directed Evolution of Androgen Receptor

Step three consists of two verification checkpoint assays: HIS3 and β-galactosidase assays (Figure 7). The *HIS3* transcription assay tests for hormone dependent or independent transcription of the AR mutants. Viable AR mutants may activate hormone-dependent or hormone-independent transcription. Colonies from the selection assay are tested for growth on agar plates with (SC-LW) and without histidine (SC-HLW). The colonies that grow on both plates activate hormone-independent transcription while those that only grow on SC-LW plates activate

hormone-dependent transcription. We are only interested in AR mutants that activated hormone-dependent transcription.

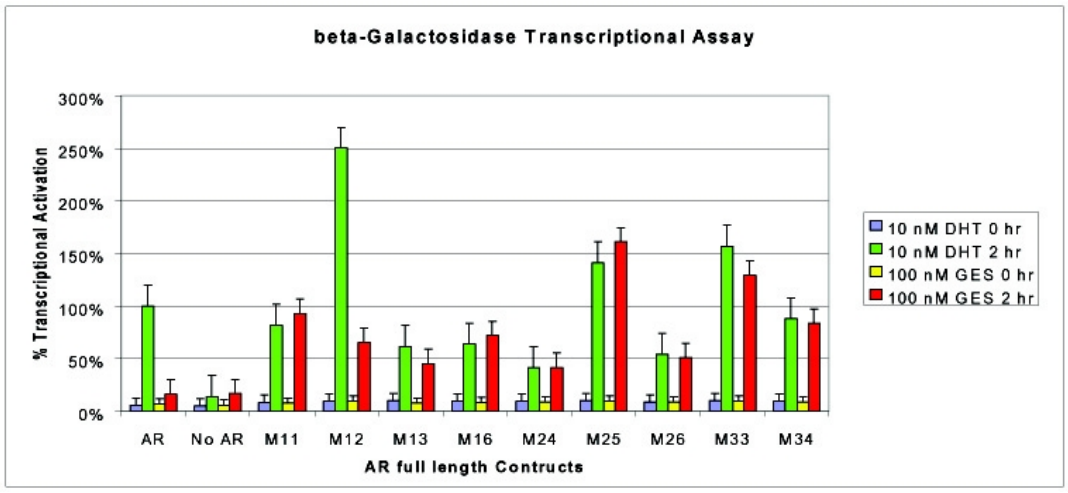


Figure 8: Transcriptional Assay Data of 9 AR mutants

The percent transcriptional activation is compared between AR wild-type and AR mutants generated in the selection assay. Nine mutants showed hormone-dependent transcription.

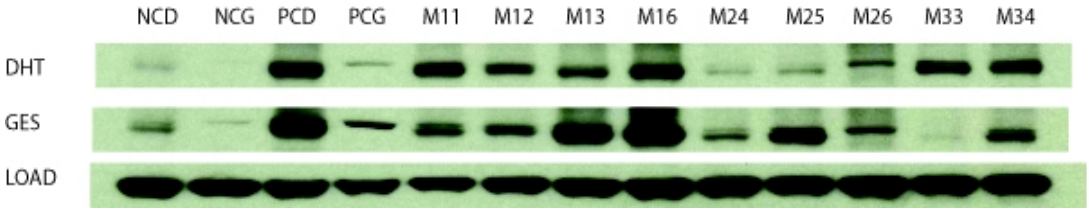


Figure 9: Protein expression levels of full-length AR constructs

Levels of protein expression of full-length AR constructs at 2 hr with 10 nM DHT and 100 nM gestrinone in β-galactosidase transcription assay. NCD; Negative control (empty vector, 10nM DHT), NCG; Negative control (empty vector, 100 nM GES),

PCD; Positive control (AR vector, 10nM DHT), PCG; Positive control (AR vector, 100 nM GES) Controls should be identical between both rows. Protein expression levels do not exactly correlate with the transcription levels seen in the yeast assay. Equivalent levels of expression are seen for M11, M12, M26, and M34. More AR expression in the presence of GES is seen for M24 and M25.

AR Mutant	DBD	Hinge	LBD
M11	E588G		
M12			G743E, A748D, P801L
M13			T860S
M16		K609R, K658R	N756D, I898V
M24		E668V	H714Q, I816N
M25			T755S
M26			S888N
M33		G627R	V730M, F856L, V901A
M34		E654V, T656S	N691S, H714Q, A735D, Q802H

Table 3: Point Mutations in AR mutants isolated from DE experiment

The AR mutant plasmids were isolated from the yeast and sequenced. The point mutations were found throughout the DNA-binding domain, hinge, and ligand-binding domain. The majority of the mutations were found in the ligand-binding domain. Upon first inspection, the mutation were mainly subtle, surface-exposed, and in the turns between helixes.

Sequence Alignments Of AR Mutations

Sequences alignments of various species of AR (human, rat, mouse, chicken, frog, japanese fish) reveal that some of the point mutations have more variability than others. E654V, K658, E668, H714, and A735 are all highly variable from species to

species. E588, K609, A748, T755, N756, Q802, and I816 are all high conserved throughout various species of AR. These point mutations may contribute more the the gain-of-function activity of the AR mutants, which contain these point mutations. Sequence and structural alignments with mammalian progesterone receptor (PR) and estrogen receptor (ER) (α and β isoforms) reveal that six AR mutant residues match the corresponding residues in PR and ER (α and β isoforms). Point mutations matching progesterone receptor are K658R, I898V, and V901A. Point mutations matching ERα are T656S, G743E, and F856L. Point mutations matching ERβ are G743E and V901A. These mutations are most likely contributing most to the gain-of-function activity seen for the corresponding AR mutants.

Mutations	***DE --> PR***	***DE --> ERα***	***DE --> ERβ***
T656S		T656S	
K658R	K658R		
G743E		G743E	G743E
F856L		F856L	
I898V	I898V		
V901A	V901A		V901A

Table 4: Analysis of Point Mutations Matching PR and ERα/β

Mutant #	Residue #	Location	Description
11	E588G	DBD	Conserved residue throughout AR species
12	G743E	Helix 5	DE to ERα and ERβ
12	A748D	Helix 5	Conserved residue throughout AR species
12	P801L	Btw H7 + H8	N-terminus of H8
13	T860S	Helix 10	Side chain interacts with F916-H917
16	K609R	Hinge	Conserved residue throughout AR species
16	K658R	Hinge	Highly variable among AR species
16	N756D	Helix 5	Conserved residue throughout Mammallian AR
16	I898V	Helix 12 (AF-2)	DE to PR
24	E668V	Hinge/H1	Highly variable among AR species
24	H714Q	Helix 3	Highly variable among AR species
24	I816N	Btw H8 + H9	Conserved residue throughout AR species and PR
25	T755S	Helix 5	Conserved residue throughout AR species
26	S888N	Btw H11 + H12	N-terminus of H12
33	G627R	Helix 1	DE to PR
33	A735D	Helix 4	Highly variable among AR species
33	F856L	Helix 10	DE to ERα
33	V901A	Helix 12	DE to PR and ERα
34	E654V	Hinge	Highly variable among AR species
34	T656S	Hinge	DE to ERα
34	N691S	Btw H1 + H3	Mutation results in H-Bond break with R774
34	H714Q	Helix 3	Highly variable among AR species
34	V730M	Helix 4	Known Potent Cancer Mutation
34	Q802H	Btw H7 + H8	Conserved residue throughout Mammallian AR

Table 5: Detailed Analysis of Point Mutations in AR Mutants

Distribution Of Point Mutations

The point mutations from the directed evolution yeast-based selection assay are distributed throughout the DBD, hinge, and LBD. Point mutations are clustered in the hinge and helices 3, 4, 5, 10, 12. Mutations in the turns are between H1/H3, H8/H9, and H11/H12. None of the mutations are in the ligand-binding pocket (Table xx).. When mapped on to the known crystal structure of AR LBD, the point mutations are main on the surface and not in contact with the ligand binding pocket (Figure 10). On the front side, point mutations on helix 3,4,5,12 can be seen. On the

left facing side, 5 mutations are lined up along a groove formed by helix 5 (T755S, N756D), the turns between helices 7 and 8 (P801L, Q802H) and the chain preceding helix 3 (N691S). This may be a protein binding interface. More experimentation is necessary. N691S is the only mutation that can be explained with molecular modeling. The N-to-S mutation causes a break in a critical hydrogen bonds with R774 on helix 6. The left side shows the mutations (I898V, V901A) on helix 12 most prominently. Interestingly, these two mutations are part of the AF-2 site. Slightly visible are I816N (between helices 8 and 9), V730M and A735D (helix 4), and S888N (between helices 11 and 12). These residues can arguable be part of the AF-2 interface. On the back side, there are two mutations T860S (helix 11) and I816N (between helices 8 and 9).

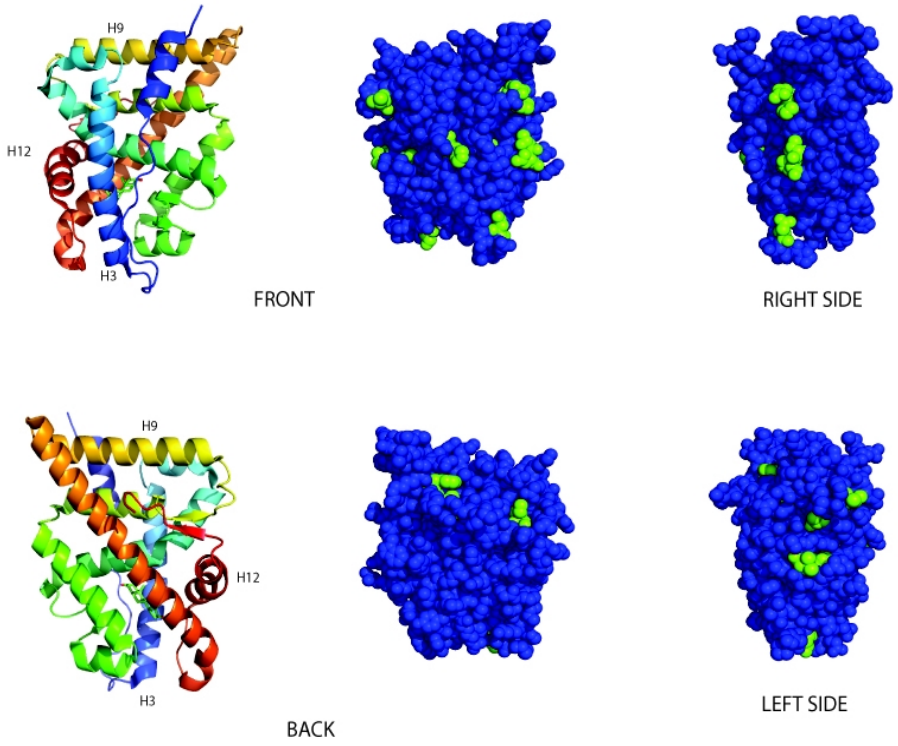


Figure 10: Structural Distribution of Directed Evolution AR Point Mutants

Future Directions

The discovery of full length AR mutants capable of transcribing with gestrinone via directed evolution is a major milestone for this project. However, further work must be done to achieve an orthogonal ligand-receptor pair of androgen receptor. Characterization of the key mutations of the 9 AR mutants with yeast and mammalian transcription assays is necessary.

First, follow-up yeast transcription assays without FKBP52 must be executed to determine the effect on the transcription levels of the AR mutants compared to wild-type. Overall, the levels of transcription should be less because FKBP52 is known to enhance the levels of transcription. If the AR mutants have a much higher level of transcription in the presence of gestrinone compared to wild-type, then the mutants have become FKBP52-independent. If the AR mutants have a much lower level of transcription in the presence of gestrinone compared to wild-type, then the mutants have become more FKBP52-dependent.

Second, each of the nine AR mutants should be deconstructed to determine which of the point mutations contribute to the observed transcription with gestrinone. To to this, each single point mutation should be tested in the yeast transcription assay to determine which of the mutation are key to the observed transcription levels for each of the AR mutants. If the point mutation contributes to the observed gestrinone-

induced transcription, there will be a measurable amount of transcription in the yeast assay. If there is not a measurable amount of transcription in the yeast assay, the point mutation does not contribute to the observed gestrinone-induced transcription. Once the key mutations are determined, the key point mutations for each original AR mutant should be inserted into the same vector using site-directed mutagenesis and tested again in the yeast assay. The levels of transcription may increase due to removal of any non-key point mutations. The selectivity may also increase for gestrinone over DHT. Also, a matrix of mutants can be created to see if there are any beneficial effects of combining point mutations that were not originally in the same AR mutant. Based on the alignments with PR, ERα, and ERβ, it would be interesting test AR mutants with only PR- directed mutations, ERα- directed mutations, and ERβ- directed mutations. Also, a focused matrix of these mutations would allow for a super-AR mutant to be found.

Stability tests should be preformed on the any constructs starting with the key point mutations. The stability of the AR mutant proteins should be similar or better that wild-type. The stability test should be performed using differential scanning fluorimetry (DSF). DSF monitors unfolding of proteins in the presence of a fluorescent dye (ex. SYPRO orange) and is typically performed by using a real-time PCR instrument (Mx3005, Stratagene). The temperature at which a protein unfolds (Tm) is measured by an increase in the fluorescence of the dye with affinity for hydrophobic parts of the protein that become exposed as the protein unfolds. DSF can

be used to detect ligand binding of a protein of interest (Niesen et. al., 2007). Indeed, most ligands stabilize proteins upon binding, causing a detectable increase in the transition temperature. The difference in Tm in the presence and absence of a tested compound is related to the binding affinity of the small molecule. Ensuring that the AR mutant protein is stable will increase the chances of finding more successful AR mutants in subsequent rounds of directed evolution.

Third, the full-length AR mutant constructs should be tested in mammalian cells to test if the AR mutants can transcribe in a more native environment. Since yeast do not have steroid nuclear receptors, the AR constructs must show robust AR protein expression and transcription of target gene (ex. luciferase) in mammalian cells. Several cell lines should be tested since it is known that full length AR does not transcribe well in some immortalized cell lines. HeLa, CV1, and DU145 cells lines will be tested first with full-length wild-type AR with 10-100 nM DHT to determine the optimal assay conditions. Western blot analysis will accompany the transcription assay to determine AR protein expression levels and to normalize the transcriptional activation data. If the AR mutants show robust transcription with gestrinone similar to wild-type AR with DHT, then the AR mutant is good. The next step is specificity of gestrinone over DHT. Ultimately, the AR mutants will have to be subjected to the directed evolution yeast assay in the presence of the novel steroids, GS1-5. For a true orthogonal ligand-receptor pair, the ligand must not activate any receptor other than

it's re-engineered receptor. Gestrinone is a known agonist of progesterone receptor and is only used as a proof-of-concept drug in this assay.

Discussion

Despite the existence of other orthogonal ligand-receptor pairs, attempting to re-engineer the androgen receptor was not as straightforward as other receptor pairs published in the literature. There was no magical residue that could be mutated to an alanine resulting in an orthogonal ligand-receptor pair. Computationally-derived AR mutants failed by only achieving, at most, 5% transcriptional activation in cell based transcription assays. Directed evolution of full-length AR in a yeast-based selection assay produced 9 AR mutants capable of activating transcription with gestrinone, an endometriosis drug that does not activate transcription with wild-type AR. At this stage, it is unknown why the point mutations alter each AR mutant to gain function with gestrinone. Additional experiments are needed. To theorize: AR may be evolving towards progesterone and estrogen receptors that is allowing AR to be activated by gestrinone. The point mutations may be stabilizing AR in an excited state that is more sensitive than wild-type AR. Given that the point mutations are dispersed throughout the DBD, hinge, and LBD, any number of factors may be contributing to the gain-of-function transcription with gestrinone.

Materials and Methods

Protein Expression And Purification

AR-LBD (residues 663–919) was expressed in *Escherichia coli* and purified to homogeneity using a modified version of previously published protocols (Hur et. al. 2004; Estébanez-Perpiñá et. al. 2005). Bacterial cell preparations were grown at ambient or lower temperatures to high optical density at 600 nm (>1.00) in 2× LB supplemented with DHT. AR-LBD protein was expressed by induction with isopropyl 1-thio-β-D-galactopyranoside for 14–16 h at 15° C before harvest and cell lysis by freeze-thawing and mild sonication. Purification involved an initial affinity chromatography step using a Ni-NTA column. Finally cation exchange chromatography with Sepharose SP afforded the purified protein.

Mutagenesis

The desired mutations was introduced by site-directed mutagenesis with the QuikChange Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA), using the conditions provided by the manufacturer.

Cell Culture and Transfection Assay

HeLa cells were maintained in Dulbecco's modified Eagle's medium H-21 4.5 g/liter glucose, containing 10% steroid depleted fetal bovine serum (Invitrogen), 2 mM glutamine, 50 units/ml penicillin, and 50 mg/ml streptomycin. For transfection, cells were collected and re-suspended in Dulbecco's phosphate-buffered saline (0.5 mL/4.5 × 107 cells) containing 0.1% dextrose, and typically 4 μg of luciferase reporter plasmid, 1 μg of AR expression vector or empty vector control, and 2 μg of pCMV-β-galactosidase. Cells were electroporated at 240 V and 960 μF, transferred to fresh media, and plated into 12-well plates. After incubation for 24 h at 37° C with androgen or vehicle, cells were collected, and pellets were lysed by addition of 150 μL of 100 mM Tris-HCl, pH 7.8, containing 0.1% Triton X-100.

For transfections with full-length AR, the reporter gene utilized the Mouse Mammary Tumor Virus promoter fused to luciferase. For transfections with *GAL-AR-LBD*, the reporter contained five *GAL4* response elements upstream of a minimal promoter. LUC and β-galactosidase activities were measured using the Luciferase Assay System (Promega) and Galacto-Light Plus β-galactosidase reporter gene assay system (Applied Biosystems), according to the manufacturer's instructions.

Partial Synthesis Of 19PT

3β-17β-androst-5-ene diol

Sodium borohydride (3.28 g, 0.0867 mol) was added in four equal portions (about 2 minutes apart) to a cold (0° C) solution of dehydroepiandrosterone (25.0g, 0.0867

mol) in methanol (870 mL). The cold bath was removed and the cloudy white mixture was stirred at room temperature for 90 minutes. The reaction mixture was cooled in an ice bath as 2N HCl (173 mL, 0.346 mol) was added dropwise. The reaction mixture was concentrated under vacuum to a wet white solid. Water (500 mL) was added and the mixture was sonicated and filtered. The collected solid was washed (100 mL) and dried in a vacuum dessicator overnight to afford the title compound as a white solid.

3β-17β-androst-5-ene diol diacetate

Acetic anhydride (19.5 mL, 0.2 mol) was added to a solution of 3β-17β-androst-5-ene diol (15.0 g, 0.05165 mol) in pyridine (200 mL) (note: the addition was mildly exothermic) then 4-dimethylamino-pyridine (0.63 g, 0.00516 mol) was added. The resulting yellow solution was stirred at room temperature for 5.5 hours then most of the solvent was removed under vacuum. The residual yellow-white sludge was partitioned between ethyl acetate (450 mL) and 1N HCl (450 mL). The organic layer was washed with 5% aqueous bicarbonate (200 mL) then dried over magnesium sulfate, filtered, and evaporated to an off-white solid. This crude product was recrystallized from hexane (500 mL) to afford the title compound as a white crystalline solid. Concentration of the mother liquor from the mother liquor from the recrystallization afforded an off-white solid which could be recrystallized to afford a second crop of product.

5α-bromo-6β-hydroxy-3β,17β-androstane diol diacetate

A solution of 70% perchloric acid (0.79 mL) in water (6.8 mL) was added to a solution of 3β,17β-androstane diol diacetate (4.17 g, 0.011 mol) in dioxane (56 mL) and water (3.4 mL) at 5° C. N-bromoacetamide (2.25 g, 0.016 mol) was added in small portions over a 20 minute period. The resulting mixture was stirred at 5° C for 30 minutes then stirred at room temperature for 30 minutes then poured into water containing 0.5 mL of 1 % sodium thiosulfate solution. The suspension was adjusted to pH 8 by addition of saturated aqueous sodium bicarbonate solution then extracted with ethyl acetate. The organic layer was washed with brine, dried over magnesium sulfate, filtered, and concentrated under vacuum to afford a white foam. The residue was combined with 0.296 g of crude product from an earlier batch and purified by recrystallization from acetone/hexane to afford the title compound as a white solid.

5α-bromo-6β,19-epoxy-3β,17β-androstane diol diacetate

Iodobenzne diacetate (1.23 g, 0.0057 mol) was added to a suspension of the product of step 1 (1.8 g, 0.0038 mol) in cyclohexane (250 mL) then iodine (0.97 g, 0.0038 mol) was added. The resulting mixture was irradiated with a 200W sun lamp for 45 minutes (note: the temperature of the mixture rose to about 80° C during this time). The reaction mixture was cooled to room temperature and poured into ice/water solution. The resulting mixture was extracted with ether. The organic layer was washed with 2% aqueous sodium thiosulfate and water then dried over magnesium

sulfate, filtered, and concentrated under vacuum. The residue was recrystallized from hexane to afford an off-white solid.

3β-17β-androst-5-ene triol 3,17-diacetate

A mixture of activated zinc dust (11.1 g, 0.17 mol; activated before use by brief treatment with aqueous HCl followed by sequential washing with water and acetone then drying under vacuum) and the product of step 2 (1.50 g, 0.0032 mol) in tetrahydrofuran (75 mL) and water (7.5 mL) was stirred at 65° C for 1 hour. The reaction mixture was cooled to room temperature and filtered. The collected solid was washed with ether then the combined filtrate was washed with water, dried over magnesium sulfate, filtered, and concentrated under vacuum to afford a pale yellow foam. The residue was recrystallized from acetone/hexane to afford the title compound as a pale yellow solid. Concentration and recrystallization of the mother liquor afforded a second crop of less pure product as a pale yellow solid.

19-oxo-3β-17β-androst-5-ene diol diacetate

Activated 4A molecular sieves (4.2 g) were added to a cold (0° C) solution of the product of step 3 (0.500 g, 0.00128 mol) and N-methylmorpholine N-oxide (NMO, 2.43 g, 0.0207 mol) in dichloromethane (10 mL). The resulting mixture was stirred at 0° C for 15 minutes then tetrapropylammonium perruthenate (0.030 g, 0.0000854 mol) was added. The resulting mixture was stirred at 0° C for 90 minutes then diluted with ether and filtered. The collected solid was washed with ether. The combined filtrate was washed sequentially with aqueous sodium sulfite, and aqueous copper

sulfate, then dried over magnesium sulfate, filtered, and concentrated under vacuum to afford a white solid. The residue was purified by flash chromatography on silica gel eluted with 95:5 dichloromethane: ethyl acetate to afford the title compound as a pale yellow solid.

Synthesis of Gestrinone Analogs

GS1

Iodomethane 1h (103 mg, 0.5 mmol, 1 equiv) is mixed with gestrinone (100 mg, 0.5 mmol, 1 equiv) in a 20 mL scintillation vial. To the mixture were added L-proline (12 mg, 0.1 mmol, 0.2 equiv), Na_2CO_3 (12 mg, 0.1 mmol, 0.2 equiv), NaN_3 (39 mg, 0.6 mmol, 1.2 equiv), sodium ascorbate (20 mg, 0.05 mmol, 0.1 equiv), 9:1 ACN/ H2O (1 mL), and $CuSO_4 \bullet 5H_2O$ (13 mg, 0.025 mmol, 0.05 equiv). The mixture was heated overnight at 65° C. Upon completion (monitored by TLC or LC-MS), the crude mixture was poured into dilute NH_4OH (30 mL; ATTENTION: this step is important, as copper azides are explosive when dry, and their traces should be removed before the product is dried) and extracted with ethyl acetate (3 X 20 mL). The organic layer was washed with brine (2 X 20 mL), dried over $MgSO_4$, and evaporated to yield the title compound as a pale yellow oil (112 mg, 94%). The pale yellow oil was further purified by flash chromatography with silica gel and 1:1 ethyl acetate/hexane to afford 20 mg (25%).

GS2

1-Iodo-ethane 1h (103 mg, 0.5 mmol, 1 equiv) is mixed with gestrinone (100 mg, 0.5 mmol, 1 equiv) in a 20 mL scintillation vial. To the mixture were added L-proline (12 mg, 0.1 mmol, 0.2 equiv), Na_2CO_3 (12 mg, 0.1 mmol, 0.2 equiv), NaN_3 (39 mg, 0.6 mmol, 1.2 equiv), sodium ascorbate (20 mg, 0.05 mmol, 0.1 equiv), 9:1 ACN/ H2O (1 mL), and $CuSO_4 \bullet 5H_2O$ (13 mg, 0.025 mmol, 0.05 equiv). The mixture was heated overnight at 65° C. Upon completion (monitored by TLC or LC-MS), the crude mixture was poured into dilute NH_4OH (30 mL; ATTENTION: this step is important, as copper azides are explosive when dry, and their traces should be removed before the product is dried) and extracted with ethyl acetate (3 X 20 mL). The organic layer was washed with brine (2 X 20 mL), dried over $MgSO_4$, and evaporated to yield the title compound as a pale yellow oil (112 mg, 94%). The pale yellow oil was further purified by flash chromatography with silica gel and 1:1 ethyl acetate/hexane to afford 25 mg (30%).

GS3

1-Iodopropane 1h (103 mg, 0.5 mmol, 1 equiv) is mixed with gestrinone (100 mg, 0.5 mmol, 1 equiv) in a 20 mL scintillation vial. To the mixture were added L-proline (12 mg, 0.1 mmol, 0.2 equiv), Na_2CO_3 (12 mg, 0.1 mmol, 0.2 equiv), NaN_3 (39 mg, 0.6 mmol, 1.2 equiv), sodium ascorbate (20 mg, 0.05 mmol, 0.1 equiv), 9:1 ACN/ H2O (1 mL), and $CuSO_4 \bullet 5H_2O$ (13 mg, 0.025 mmol, 0.05 equiv). The mixture was heated overnight at 65° C. Upon completion (monitored by TLC or LC-MS), the crude

mixture was poured into dilute NH_4OH (30 mL; ATTENTION: this step is important, as copper azides are explosive when dry, and their traces should be removed before the product is dried) and extracted with ethyl acetate (3 X 20 mL). The organic layer was washed with brine (2 X 20 mL), dried over $MgSO_4$, and evaporated to yield the title compound as a pale yellow oil (112 mg, 94%). The pale yellow oil was further purified by flash chromatography with silica gel and 1:1 ethyl acetate/hexane to afford 15 mg (20%).

GS4

1-Iodobutane (103 mg, 0.5 mmol, 1 equiv) is mixed with gestrinone (100 mg, 0.5 mmol, 1 equiv) in a 20 mL scintillation vial. To the mixture were added L-proline (12 mg, 0.1 mmol, 0.2 equiv), Na_2CO_3 (12 mg, 0.1 mmol, 0.2 equiv), NaN_3 (39 mg, 0.6 mmol, 1.2 equiv), sodium ascorbate (20 mg, 0.05 mmol, 0.1 equiv), 9:1 ACN/ H_2O (1 mL), and $CuSO_4 \bullet 5H_2O$ (13 mg, 0.025 mmol, 0.05 equiv). The mixture was heated overnight at 65° C. Upon completion (monitored by TLC or LC-MS), the crude mixture was poured into dilute NH_4OH (30 mL; ATTENTION: this step is important, as copper azides are explosive when dry, and their traces should be removed before the product is dried) and extracted with ethyl acetate (3 X 20 mL). The organic layer was washed with brine (2 X 20 mL), dried over $MgSO_4$, and evaporated to yield the title compound as a pale yellow oil (112 mg, 94%). The pale yellow oil was further purified by flash chromatography with silica gel and 1:1 ethyl acetate/hexane to afford 20 mg (25%).

GS5

1-Iodo-2-methylpropane (103 mg, 0.5 mmol, 1 equiv) is mixed with gestrinone (100 mg, 0.5 mmol, 1 equiv) in a 20 mL scintillation vial. To the mixture were added L-proline (12 mg, 0.1 mmol, 0.2 equiv), Na_2CO_3 (12 mg, 0.1 mmol, 0.2 equiv), NaN_3 (39 mg, 0.6 mmol, 1.2 equiv), sodium ascorbate (20 mg, 0.05 mmol, 0.1 equiv), 9:1 ACN/ H2O (1 mL), and $CuSO_4 \bullet 5H_2O$ (13 mg, 0.025 mmol, 0.05 equiv). The mixture was heated overnight at 65° C. Upon completion (monitored by TLC or LC-MS), the crude mixture was poured into dilute NH_4OH (30 mL: this step is important, as copper azides are explosive when dry, and their traces should be removed before the product is dried) and extracted with ethyl acetate (3 X 20 mL). The organic layer was washed with brine (2 X 20 mL), dried over $MgSO_4$, and evaporated to yield the title compound as a pale yellow oil (112 mg, 94%). The pale yellow oil was further purified by flash chromatography with silica gel and 1:1 ethyl acetate/hexane to afford 15 mg (20%).

Yeast hormone induction assays, methods, plasmids, and strains

The yeast hormone induction assay protocol, *Saccharomyces cerevisiae* strain DSY-1513 and the plasmids (AR full length, AR (559-919), and FKBP52 expression plasmids) have been previously described (Riggs et. al., 2007). Mutations were introduced by site-directed mutagenesis (QuikChange II; Stratagene, San Diego, CA) into the wild-type human AR gene cloned into p425GPD. Potentiation by these mutant ARs was measured in strain DSY-1513 transformed with the FKBP52 yeast

plasmid (p424GPD_FKBP52). Dihydrotestosterone (DHT) (10 nM) was used in the AR hormone induction assays.

Directed Evolution Selection Assay and Analysis

The mutant AR library was made using error-prone PCR (GeneMorphII; Stratagene, CA) using the manufacturer's recommended conditions for high-frequency mutagenesis (5 ng target DNA per reaction). The template used was p424GPD-hAR, and the primer binding sites were approximately 100 bases outside of the gene borders. The PCR product (3 μg, purified by agarose gel electrophoresis) and p424GPD vector (1 μg, linearized with HindIII and XhoI) were co-transformed into the selection strain DSY-1513 (*Ta ura3-52 lys2-801 ade2-101 trp1-Δ63 his3- Δ200 leu2-Δ1 pdr5::GT3Z his3::GT3H*; gift of Natasha Kralli, Scripps Institute) containing the plasmid p425GPD-FKBP52. AR potentiation mutants were selected on plates containing synthetic complete medium lacking tryptophan, leucine, and histidine (SC-HLW) supplemented with 100 nM gestrinone (or GS1-5) and 5 mM 3-amino-1,2,4-triazole. The colonies appearing after about 1 week of incubation at 30° C were purified on SC-WL plates. The mutant phenotype (potentiation of hormone signaling) was confirmed by assaying hormone-dependent expression of the β-galactosidase reporter in the SC-HLW selection strain. Plasmids of AR mutants that showed increases in both hormone-dependent growth and β-galactosidase expression were isolated by extracting the mutated AR plasmids from yeast lysates (Wizard Plus SV Minipreps, Promega Corporation, Madison, WI). Forward and

reverse primers were used in the DNA sequencing of the mutated region to confirm the single point mutations.

Western Immunoblots

To confirm equivalent expression of each AR form, Western immunoblots were performed on extracts from yeast cells used in the transcriptional activation assays. Yeast extracts were prepared with glass beads as previously described (Riggs et. al., 2007)). The following mouse monoclonal antibodies were used: anti-AR (epitope in NTD) (Santa Cruz Biotech. #SC7305). To confirm sample loading, antibody against L3 ribosomal protein (yeast lysates) (Riggs et. al. 2007) were used.

References

Abate-Shen C, Shen MM (2000) Molecular genetics of prostate cancer. Genes Dev 14: 2410-2434.

Alberts B, Johnson A, Lewis J, Raff M, Roberts, K, Walter P (2002) Molecular Biology of the Cell. 4th Edition. Garland Science

Aranda A, Pascual A (2001) Nuclear hormone receptors and gene expression. Physiol Rev. 81:1269-1304.

Arnold JT & Isaacs JT (2002) Mechanisms involved in the progression of androgen-independent prostate cancers: it is not only the cancer cell's fault. Endocr Relat Cancer 9:61-73.

Berg J, (1989) DNA-binding specificity of steroid receptors. Cell 57: 1065-1068.

Blanchet, J., Lin, S.-X., and Zhorov, B.S. (2005) Mapping of steroids bindingto 17b-hydroxysteroid dehydrogenase type 1 using Monte Carlo energy minimization reveals alternative binding modes. Biochemistry 44: 7218–7227.

Blizzard, TA et. al., (2006a) Estrogen Receptor Modulators WO_200505197141

Blizzard TA et. al., (2006b) Androstenediol analogs as ERβ-selective SERMS Bioorg. Med Chem Lett 16: 834-838.

Bolander FF (2004) Molecular Endocrinology 3rd Edition, Elsevier Academic Press

Brannigan JA, Wilkinson AJ (2002) Protein engineering 20 years on. Nature Reviews 3:964-970.

Brinkmann AO, Faber PW, van Rooij HJC, et al, (1989) The human androgen receptor: domain structure, genomic organization, and regulation of expression. J Steroid Biochem 34: 307-310.

Brinkmann AO (2001) Molecular basis of androgen insensitivity. Mol Cell Endocrinol 179: 105-109.

Brown, C. J., Goss, S. J., et. al. (1989)Androgen receptor locus on the human X chromosome: regional localization to Xq11-12 and description of a DNA polymorphism. Am J Hum Genet 44:264-9.

Buskirk, A. R.; Liu, D. R. (2005) Creating small-molecule-dependent switches to modulate biological functions. Chem. Biol. 12, 151-161.

Chang, C. S., Kokontis, J., et. al. (1988) Molecular cloning of human and rat complementary DNA encoding androgen receptors. Science 240:324-6

Catlin D.H., Sekera M.H., Ahrens B.D., Starcevic B., Chang Y.C., Hatton C.K. (2004) Tetrahydrogestrinone: Discovery, synthesis, and detection in urine Rapid Commun. Mass Spectrom. 18 1245–1249.

Chen Z, Katzenellenbogen BS, Katzenellenbogen JA, Zhao H (2004) Directed Evolution of Human Estrogen Receptor Variants with Significantly Enhanced Androgen Specificity and Affinity J Bio Chem 279:33855-33864.

Chockalingam K, Chen Z, Katzenellenbogen JA, Zhao (2005) Directed Evolution of Specific receptor-ligand pairs for use in the creation of gene switches 102: 5691-5696.

Dawood M.Y., Obasiolu C.W., Ramos J., Khan-Dawood F.S. (1997) Clinical, endocrine, and metabolic effects of two doses of gestrinone in treatment of pelvic endometriosis Am. J. Obstet. Gynecol. 176 387–394.

Death A.K., McGrath K.C., Kazlauskas R., Handelsman D.J. (2004) Tetrahydrogestrinone is a potent androgen and progestin J. Clin. Endocrinol. Metab. 89 2498–2500.

Dubbink HJ, Hersmus R, Verma CS, van der Korput HA, Berrevoets CA, van Tol J, Ziel-van der Made AC, Brinkmann AO, Pike AC, Trapman J (2004). "Distinct recognition modes of FXXLF and LXXLL motifs by the androgen receptor". Mol. Endocrinol. 18: 2132–50.

Escriva H, Bertrand S, Laudet V (2004) The evolution of the nuclear receptor superfamily. Essays Biochem. 40:11-26.

Estébanez-Perpiñá E., Moore J.M., Mar E., Delgado-Rodrigues E., Nguyen P., Baxter J.D., Buehrer B.M., Webb P., Fletterick R.J., Guy R.K. (2005) The molecular mechanisms of coactivator utilization in ligand-dependent transactivation by the androgen receptor J. Biol. Chem. 280 8060–8068.

Evans RM, (1988) The steroid and thyroid hormone receptor superfamily. Science 240: 889-895.

Feldman AK, Colasson B, Folkin VV (2004) One-pot Synthesis of 1,4-Disubstituted 1,2,3-Triazoles from In Situ Generated Azides Org Lett 6:3897-3899

Freedman LP, (1992) Anatomy of the steroid receptor zinc finger region. Endocr Rev 13: 129-145.

Gelmann EP (2002) Molecular biology of the androgen receptor. J Clin Oncol 20: 3001-3015.

Georget V, Lobaccaro JM, Terouanne B, Mangeat P, Nicolas JC, Sultan C (1997) Trafficking of the androgen receptor in living cells with fused green fluorescent protein-androgen receptor. Mol Cell Endocrinol 129: 17-29.

Georget V, Bourguet W, Lumbroso S, Makni S, Sultan C, Nicolas JC (2006) Glutamic acid 709 substitutions highlight the importance of the interaction between androgen receptor helices H3 and H12 for androgen and antiandrogen actions. Mol Endocrinol. 20:724-734.

Glass CK, Rosenfeld MG (2000) The coregulator exchange in transcriptional functions of nuclear receptors. Genes Dev 14: 121-141

Hashimoto, A, Shi Y, Drake K, Koh, JT (2005) Design and synthesis of complementing ligands for mutant thyroid hormone receptor TRbeta(R320H): a tailor-made approach toward the treatment of resistance to thyroid hormone. Bioorg. Med. Chem. 13, 3627-3639.

He B, Gampe Jr RT, Kole, A.J., Hnat, A.T., Stanley, T.B., An, G.,Stewart, E.L., Kalman, R.I., Minges, J.T., and Wilson, E.M. (2004) Structural basis for androgen receptor interdomain and coactivator interactions suggests a transition in nuclear receptor activation function dominance.Mol. Cell 16: 425–438.

Heinlein CA, Chang C (2002) Androgen receptor (AR) coregulators: an overview. Endocr Rev. 23:175-200.

Hiort O, Holterhus PM, (2003) Androgen insensitivity and male infertility. Int J Androl 26: 16-20.

Hollenberg SM, Weinberger C, Ong ES, Cerelli G, Oro A, Lebo R, Thompson EB, Rosenfeld MG, Evans RM (1985) Primary structure and expression of a functional human glucocorticoid receptor cDNA. Nature. 318:635-641

Hur, E., Pfaff, S.J., Payne, E.S., Gron, H., Buehrer, B.M., and Fletterick, R.J. (2004) Recognition and accommodation at the androgen receptor coactivator binding interface. PLoS Biol. 2: E274.

Hwang, Y. W.; Miller, D. L. (1987) A mutation that alters the nucleotide specificity of elongation factor Tu, a GTP regulatory protein. J. Biol. Chem. 1987, 262, 13081-13085.

Jackel C, Kast P, Hilvert D (2008) Protein Design by Directed Evolution Annu. Rev. Biophys. 37:153-173.

Jarow JP, Wright WW, Brown TR, Yan X, Zirkin BR (2005) Bioactivity of androgens within the testes and serum of normal men .J Androl. 26:343-348.

Jenster G, van der Korput HAGM, van Vroonhooven C, van der Kwast TH, Trapman J, Brinkmann AO, (1991) Domains of the human androgen receptor involved in steroid binding, transcriptional activation and subcellular localization. Mol Endocrinol 5: 1396-1404.

Keay J, Thornton JW. Hormone-activated estrogen receptors in annelid invertebrates: implications for evolution and endocrine disruption. Endocrinology 150:1731-1738, 2009.

Koh, J. T.; Biggins, J. B. (2005)Ligand-receptor engineering and its application towards the complementation of genetic disease and target identification. Curr. Top. Med. Chem. 2005, 5, 413-420.

Kuhlman, Brian; Dantas, Gautam; Ireton, Gregory C.; Varani, Gabriele; Stoddard, Barry L. & Baker, David (2003), "Design of a Novel Globular Protein Fold with Atomic-Level Accuracy", Science 302 (5649): 1364–1368.

Kuiper GCJM, Faber PW, van Rooij HJC, et al, (1989) Structural organization of the human androgen receptor gene. J Mol Endocrinol 2: R1-R4.

La Spada AR, Wilson EM, Lubahn DB, Harding AE, Fischbeck KH, (1991) Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. Nature 352: 77-79.

Labrie F., Luu-The V., Calvo E., Martel C., Cloutier J., Gauthier S., Belleau P., Morissette J., Levesque M.H., Labrie C. (2005) Tetrahydrogestrinone induces a genomic signature typical of a potent anabolic steroid J. Endocrinol. 184:427–433.

Langley E, Zhou ZX, Wilson EM (1995) Evidence for an antiparallel orientation of the ligand-activated human androgen receptor dimer. J Biol Chem. 270:29983-29990.

Lee JW, Lee YC, Na SY, Jung DJ, Lee SK (2001) Transcriptional coregulators of the nuclear receptor superfamily: coactivators and corepressors. Cell Mol Life Sci. 58:289-297.

Lin, Q.; Jiang, F.; Schultz, P. G.; Gray, N. S. Design of allele-specific protein methyltransferase inhibitors. J. Am. Chem. Soc. 2001, 123, 11608-11613.

Liu, Y.; Shah, K.; Yang, F.; Witucki, L.; Shokat, K. M. (1998) Engineering Src family protein kinases with unnatural nucleotide specificity. Chem. Biol. 1998, 5, 91-101.

Looger, Loren L.; Dwyer, Mary A.; Smith, James J. & Hellinga, Homme W. (2003), "Computational design of receptor and sensor proteins with novel functions", Nature 423 (6936): 185–190.

Lubahn DB, Joseph DR, Sar M, Tan J, Higgs HN, Larson RE, French FS, Wilson EM (1988) The human androgen receptor: complementary deoxyribonucleic acid cloning, sequence analysis and gene expression in prostate. Mol Endocrinol 2:1265-1275

Lubahn DB, Brown TR, Simantal JA, et al, (1989) Sequence of the intron/exon junctions of the coding region of the human androgen receptor gene and identification of a point mutation in a family with complete androgen insensitivity. Proc Natl Acad Sci USA 86: 9534-9538.

Mangelsdorf DJ, Thummel C, Beato M, Herrlich P, Schutz G, Umesono K, Blumberg B, Kastner P, Mark M, Chambon P, Evans RM (1995) The nuclear receptor superfamily: the second decade. Cell. 83:835-9

Marcelli M, Tilley WD, Wilson CM, Griffin JE, Wilson JD, McPhaul MJ, (1990) Definition of the human androgen receptor gene structure permits the identification of mutations that cause androgen resistance: premature termination of the receptor protein at amino acid residue 588 causes complete androgen resistance. Mol Endocrinol 4: 1105-1116.

Margolis RN, Evans RM, O'Malley BW (2005) NURSA Atlas Consortium. The Nuclear Receptor Signaling Atlas: development of a functional atlas of nuclear receptors. Mol Endocrinol. 19:2433-2436.

Matias PM, Donner P, Coelho R, Thomaz M, Peixoto C, Macedo S, Otto N, Joschko S, Scholz P, Wegg A, Basler S, Schafer M, Egner U, Carrondo MA (2000) Structural evidence for ligand specificity in the binding domain of the human androgen receptor. Implications for pathogenic gene mutations. J Biol Chem. 275:26164-26171.

Matias, P.M., Carrondo, M.A., Coelho, R., Thomaz, M., Zhao, X.Y., Wegg, A., Crusius, K., Egner, U., and Donner, P. 2002. Structural basis for the glucocorticoid response in a mutant human androgen receptor (AR(ccr)) derived from an androgen-independent prostate cancer. J. Med. Chem. 45:1439–1446.

McKenna NJ, Lanz RB, O'Malley BW (1999a) Nuclear receptor coregulators: cellular and molecular biology. Endocr Rev. 20:321-344.

McKenna NJ, Xu J, Nawaz Z, Tsai SY, Tsai MJ, O'Malley BW (1999b) Nuclear receptor coactivators: multiple enzymes, multiple complexes, multiple functions. J Steroid Biochem Mol Biol. 69:3-12.

McKenna NJ, O'Malley BW (2002) Combinatorial control of gene expression by nuclear receptors and coregulators. Cell. 108:465-474.

McKenna NJ, O'Malley BW (2002) Minireview: nuclear receptor coactivators--an update. Endocrinology. 143:2461- 2465.

McPhaul MJ, Griffin JE, (1999) Male pseudohermaphroditism caused by mutations of the human androgen receptor. J Clin Endocrinol Metab 84: 3435-3441.

McPhaul MJ (1999) Molecular defects of the androgen receptor. J Steroid Biochem Mol Biol. 69:315-322.

McPhaul MJ (2002) Molecular defects of the androgen receptor. Recent Prog Horm Res. 57:181-194.

Meehan KL, Sadar MD, (2003) Androgens and androgen receptor in prostate and ovarian malignancies. Front Biosci 8: d780-800.

Sack, J.S., Kish, K.F., Wang, C., Attar, R.M., Kiefer, S.E., An, Y., Wu, G.Y., Scheffler, J.E., Salvati, M.E., Krystek Jr., S.R., et. al. (2001) Crystallographic structures of the ligand-binding domains of the androgen receptor and its T877A mutant complexed with the natural agonist dihydrotestosterone. Proc. Natl. Acad. Sci. 98: 4904–4909.

Sultan C, Savage MO (1998) In Grossman A (ed), Clinical Endocrinology, Blackwell Science, Oxford; pp, 795-809.

Nagy L, Schwabe JW (2004) Mechanism of the nuclear receptor molecular switch. Trends Biochem Sci. 29:317-324.

Nef S, Parada LF (2000) Hormones in male sexual development. Genes Dev. 14:3075-3086.

Niculescu-Duvaz, I.; Friedlos, F.; Niculescu-Duvaz, D.; Davies, L.; Springer, C. J. (1999) Prodrugs for antibody- and gene-directed enzyme prodrug therapies (ADEPT and GDEPT). Anticancer Drug Des. 14: 517-538.

Niesen FH et. al. (2007) The use of DSF to detect ligand interactions that promote protein stability. Nature Protocols 2: 2212-2221.

Novac N, Heinzel T (2004) Nuclear receptors: overview and classification. Curr Drug Targets Inflamm Allergy. 3:335-346.

Palvimo JJ, Kallio PJ, Ikonen T, Mehto M, Janne OA, (1993) Dominant negative regulation of trans-activation by the rat androgen receptor: roles of the N-terminal domain and heterodimer formation. Mol Endocrinol 7: 1399-1407.

Parkin DM, Bray F, Ferlay J, Pisani P (2002) Global cancer statistics, CA Cancer J Clin. 55:74-108.

Pereira de Jésus-Tran K, Côté PL, Cantin L, Blanchet J, Lab rie F, Breton R (2006) Comparison of crystal structures of human androgen receptor ligand-binding domain complexed with various agonists reveals molecular determinants responsible for binding affinity. Protein Sci 15:987–999.

Pratt WB, Toft DO (1997) Steroid receptor interactions with heat shock protein and immunophilin chaperones. Endocr Rev. 18:306-360.

Pratt WB, Toft DO (2003) Regulation of signaling protein function and trafficking by the hsp90/hsp70-based chaperone machinery. Exp Biol Med. 228:111-133.

Pratt WB, Galigniana MD, Morishima Y, Murphy PJ (2004) Role of molecular chaperones in steroid receptor action. Essays Biochem. 40:41-58.

Quigley CA, De Bellis A, Marschke KB, el-Awady MK, Wilson EM, French FS, (1995) Androgen receptor defects: historical, clinical, and molecular perspectives. Endocr Rev 16: 271-321.

Riggs DL, Cox MB, Tardif HL, Hessling HL, Buchner J, Smith DF (2007) Noncatalytic Role of the FKBP52 Peptidyl-Prolyl Isomerase Domain in the Regulation of Steroid Hormone Signaling J Bio Chem 27: 8658-8669.

Robinson-Rechavi M, Escriva Garcia H, Laudet V (2003a) The nuclear receptor superfamily. J Cell Sci. 116:585-586

Robinson-Rechavi M, Laudet V (2003b) Bioinformatics of nuclear receptors. Methods Enzymol. 364:95-118

Rundlett SE, Wu X-P, Miesfeld RL, (1990) Functional characterizations of the androgen receptor confirm that the molecular basis of androgen action is transcriptional regulation. Mol Endocrinol 4: 708-714.

Shi, Y.; Koh, J.T. (2002) Functionally orthogonal ligand-receptor pairs for the selective regulation of gene expression generated by manipulation of charged residues at the ligand-receptor interface of ER alpha and ER beta. J. Am. Chem. Soc. 124, 6921-6928.

Shang Y, Myers M, Brown M (2002) Formation of the androgen receptor transcription complex. Mol Cell. 9:601-610.

Shen HC, Coetzee GA (2005) The androgen receptor: unlocking the secrets of its unique transactivation domain. Vitam Horm. 71:301-139.

Simental JA, Sar M, Lane MV, French FS, Wilson EM, (1991) Transcription activation and nuclear targeting signals of the human androgen receptor. J Biol Chem 266: 510-518.

Strausberg, R. L.; Schreiber, S. L. From knowing to controlling: a path from genomics to drugs using small molecule probes. Science 2003, 300, 294-295.

Sultan C, Paris F, Terouanne B, et al, 2001 Disorders linked to insufficient androgen action in male children. Hum Reprod Update 7: 314-322.

Tedesco, R.; Thomas, J. A.; Katzenellenbogen, B. S.; Katzenellenbogen, J. A. The estrogen receptor: a structure-based approach to the design of new specific hormone-receptor combinations. Chem. Biol. 2001, 8, 277-287.

Thornton JW, Kelley DB (1998) Evolution of the androgen receptor: structure-function implications. Bioessays. 20:860-869.

Tilley WD, Marcelli M, Wilson JD, McPhaul MJ (1989) Characterization and expression of a cDNA encoding the human androgen receptor. Proc Natl Acad Sci 86:327-331

Trapman J, Klaassen P, Kuiper GG, van der Korput JA, Faber PW, van Rooij HC, Geurts van Kessel A, Voorhorst MM, Mulder E, Brinkmann AO (1988) Cloning, structure and expression of a cDNA encoding the human androgen receptor. Biochem Biophys Res Commun 153:241-248.

Venter JC et. al., (2001) The sequence of the human genome. Science. 291:1304-1351

Warne GL, Kanumakala SH, (2002) Molecular endocrinology of sex differentiation. Semin Reprod Med 20:169-180.

Wurtz, J.M., Bourguet, W., Renaud, J.P., Vivat, V., Chambon, P., Moras, D., and Gronemeyer, H. (1996) A canonical structure for the ligand-binding domainof nuclear receptors. Nat. Struct. Biol. 3:87–94.

Zhou Z-X, Sar M, Simental JA, Lane MV, Wilson EM, (1994) A ligand-dependent bipartite nuclear targeting signal in the human andogen receptor. J Biol Chem 269: 13115-13123.

UCSF Library Release

Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

Please sign the following statement:

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.

Leslie Ann Cruz Jan 10, 2011

Author Signature Date