THERMODYNAMIC AND KINETIC ANALYSES OF ANDROGEN AND GLUCOCORTICOID RECEPTOR-
PROMOTER INTERACTIONS SUGGEST PHYLOGENETIC PARTITIONING OF ENERGETICS AND THE
PRESENCE OF FUNCTIONALLY DISTINCT DNA-BOUND STATE

by

ROLANDO W. DE ANGELIS

B.S. Indiana University, Bloomington, 2002

A thesis submitted to the
Faculty of the Graduate School of the
University of Colorado in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Pharmaceutical Sciences
2015
This thesis for the Doctor of Philosophy degree by

Rolando W. De Angelis

has been approved for the

Pharmaceutical Sciences Program

by

Tom J. Anchordoquy, Chair

David L. Bain, Advisor

Krishna M.G. Mallela

James R. Lambert

Theodore W. Randolph

Date: 12/18/2015
The fundamental principles governing transcriptional activation in higher eukaryotes are not well understood. As a step towards defining these principles, we seek a quantitative understanding of transcription factor interactions. To this end, we use thermodynamics coupled with statistical thermodynamic modeling and transient state kinetics to elucidate mechanisms of transcription factor interactions. As a model system, our lab studies the steroid receptor family of ligand-activated transcription factors. This family is comprised of two subgroups. Subgroup 3A consists of the two estrogen receptor isoforms (ER-α and ER-β), receptors that are closely related to an ancient ER-like steroid receptor ancestor. Subgroup 3C consists of more distantly related proteins: androgen receptor (AR), glucocorticoid receptor (GR), mineralocorticoid receptor (MR), and the two progesterone receptor isoforms (PR-A and PR-B). Here, we investigated the self-assembly and promoter energetics of wild type AR and a point mutation associated with prostate cancer progression, T877A. I first found that the AR constructs show no evidence of dimerization. Although this is in contrast to the literature, it is in-line with our previous studies on the other subgroup 3C members. I also found that the AR constructs display strong intersite cooperativity at a two-site promoter. Again, this is consistent with our studies on PR-B and GR, two other subgroup 3C members. Taken together, these results are in contrast to the more distantly related receptor ER-α. Thus, the steroid receptors partition their self-association and promoter assembly energetics in-line with their phylogenetic divergence. In
addition to thermodynamic studies, the dissociation kinetics of GR from a panel of six response elements and the mouse mammary tumor virus promoter were examined. In contrast to the literature, GR dissociates from its response elements on the timescale of seconds to tens of seconds, indicating that receptor-DNA interactions are primarily responsible for transient behavior. Also, GR dissociation is biphasic in character, suggesting receptor interactions with DNA are more complex than a rigid body interaction. Direct fitting of kinetic data indicates that GR dissociation involves an isomerization between two distinct GR-DNA states, suggesting two functionally distinct states.

The form and content of this abstract are approved. I recommend its publication.

Approved: David L. Bain
For Katy and Levi
ACKNOWLEDGEMENTS

Over the past six and a half years I have received encouragement and support from a number of individuals. My advisor, David Bain, has challenged me to become a critically thinking scientist and has made this journey thoughtful and rewarding. I also thank my dissertation committee of Jim Lambert, Tom Anchordoquy, Krishna Mallela, and Ted Randolph for their time and thoughtful discussions. Finally, this work would not have been possible without the support of former lab members of the laboratories of David Bain and N. Karl Maluf: Qin Yang, Keith Connaghan, Michael Miura, Teng-Chieh (Jay) Yang, Amie Moody, Nicole Fazio, and Jim Robblee.
# TABLE OF CONTENTS

CHAPTER

I. INTRODUCTION ................................................................................................................................. 1

  The Steroid Receptors: A Homologous Family of Ligand-Activated Transcription Factors ................................................................. 1

  How Do the Steroid Receptors Perform Their Specific Functions? .................................................. 3

  Overview ............................................................................................................................................ 18

II. DISSECTION OF ANDROGEN RECEPTOR-PROMOTER INTERACTIONS: STEROID RECEPTORS PARTITION THEIR INTERACTION ENERGETICS IN PARALLEL WITH THEIR PHYLOGENETIC DIVERGENCE ......................................................... 19

  Summary ........................................................................................................................................... 19

  Introduction ...................................................................................................................................... 20

  Results ............................................................................................................................................. 24

  Discussion ....................................................................................................................................... 34

  Materials and Methods .................................................................................................................. 41

III. GR-PROMOTER DISSOCIATION KINETICS MEASURED IN VITRO REVEAL EXCHANGE ON THE SECOND TIMESCALE ................................................................. 47

  Summary ........................................................................................................................................... 47

  Introduction ...................................................................................................................................... 48

  Results ............................................................................................................................................. 50

  Discussion ....................................................................................................................................... 61

  Materials and Methods .................................................................................................................. 66

IV. CONCLUSIONS ................................................................................................................................. 70

REFERENCES ......................................................................................................................................... 82

APPENDIX

A. EVALUATION OF BIPHASIC DISSOCIATION KINETICS AND RESOLUTION OF UNIQUE PARAMETERS FROM A MINIMAL KINETIC MODEL ................................................................. 91
LIST OF TABLES

TABLE

2.1 Wild type AR and T877A hydrodynamic properties .................................................. 31

2.2 Microstate energetics of wt AR and T877A assembly at the HRE2 promoter .......... 34

3.1 Observed rate constants and relative amplitudes for biphasic dissociation of GR from the TAT3 response element .............................................. 55

3.2 Observed rate constants for biphasic dissociation of GR from six individual response elements .................................................................................. 58
# LIST OF FIGURES

## FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Phylogenetic tree of the steroid receptor family depicting the family’s divergence</td>
<td>2</td>
</tr>
<tr>
<td>1.2</td>
<td>Representative schematic of a dimer binding pathway at the HRE&lt;sub&gt;2&lt;/sub&gt; promoter</td>
<td>11</td>
</tr>
<tr>
<td>2.1</td>
<td>Steroid receptor phylogenetic tree and promoter binding energetics</td>
<td>21</td>
</tr>
<tr>
<td>2.2</td>
<td>Purification of full-length human AR and characterization using sedimentation velocity and native-PAGE</td>
<td>25</td>
</tr>
<tr>
<td>2.3</td>
<td>Sedimentation equilibrium analysis of wt AR and T877A in 500 mM NaCl</td>
<td>28</td>
</tr>
<tr>
<td>2.4</td>
<td>Sedimentation velocity analysis of wt AR and T877A in 100 mM NaCl</td>
<td>30</td>
</tr>
<tr>
<td>2.5</td>
<td>Sedimentation equilibrium analysis of T877A in 100 mM NaCl</td>
<td>31</td>
</tr>
<tr>
<td>2.6</td>
<td>Quantitative DNase footprint titrations and individual site-binding isotherms of wt AR and T877A assembly at the HRE&lt;sub&gt;2&lt;/sub&gt; promoter</td>
<td>33</td>
</tr>
<tr>
<td>2.7</td>
<td>Microstate energetics of steroid receptor dimer assembly at the HRE&lt;sub&gt;2&lt;/sub&gt; promoter</td>
<td>38</td>
</tr>
<tr>
<td>2.8</td>
<td>Probabilities of receptor dimer formation and fully-ligated HRE2 promoter microspecies under standard state conditions</td>
<td>39</td>
</tr>
<tr>
<td>3.1</td>
<td>Schematics for six response elements and the mouse mammary tumor virus (MMTV) promoter</td>
<td>49</td>
</tr>
<tr>
<td>3.2</td>
<td>Purification and biophysical characterization of full-length human GR</td>
<td>51</td>
</tr>
<tr>
<td>3.3</td>
<td>Kinetic footprint and decay curve for GR dissociation from the TAT&lt;sub&gt;3&lt;/sub&gt; response element</td>
<td>53</td>
</tr>
<tr>
<td>3.4</td>
<td>Relative amplitudes and observed rate constants for concentration-dependent GR dissociation from the TAT&lt;sub&gt;3&lt;/sub&gt; response element</td>
<td>54</td>
</tr>
<tr>
<td>3.5</td>
<td>Decay curves for GR dissociation from five response elements</td>
<td>56</td>
</tr>
<tr>
<td>3.6</td>
<td>Observed rate constants for GR dissociation from six response elements</td>
<td>57</td>
</tr>
<tr>
<td>3.7</td>
<td>GR dissociation from the MMTV promoter and individual decay curves</td>
<td>59</td>
</tr>
</tbody>
</table>
3.8 Schematic and decay curve for dissociation of the GR/ER chimera from the TAT₃ sequence ................................................................. 60

3.9 Global analysis of GR-TAT₃ decay curves using a minimal kinetic model ............... 65

A.1 Error surfaces corresponding to the three parameters of the minimal kinetic model ................................................................. 95
The regulated expression of genes is essential for life. In higher eukaryotes, genes are regulated to control the activity state of a cell, helping to ensure the appropriate function and development of an organism as a whole. Transcriptional activation has been credited as playing a central role in gene regulation. Transcriptional activation is mediated by transcription factors, modular proteins that recognize specific DNA sequences and recruit co-activating proteins and the basal transcriptional machinery to initiate transcription. Biochemical studies have investigated how the basal transcriptional machinery initiates transcription and have defined key factors involved in this highly orchestrated process. However, a quantitative framework by which transcription factors initiate transcription has not been established. To establish a quantitative framework for transcription factor function, the interactions between the transcription factor and its DNA binding sites must be quantitatively investigated to establish the “rules and states” of the system. In doing so, our understanding of transcriptional activation will transition from a phenomenological understanding to a quantitative understanding with predictive power.

The Steroid Receptors: A Homologous Family of Ligand-Activated Transcription Factors

As a model system of transcriptional activation in higher eukaryotes, our lab studies the steroid receptor family of ligand-activated transcription factors. Members of this family include the androgen receptor (AR), glucocorticoid receptor (GR), mineralocorticoid receptor (MR), and two isoforms of the progesterone receptor (PR-A and PR-B) and the estrogen receptor (ER-α and ER-β). The receptors are involved in critical processes such as development, homeostasis, and
pathological disorders such as breast and prostate cancers\textsuperscript{4,5}. Phylogenetic studies demonstrate that the receptors descended from a common ER-like ancestor\textsuperscript{6,7}. Following two genome expansion events, the receptors split into two distinct subgroups. Subgroup 3A is comprised of ER-\(\alpha\) and ER-\(\beta\) and subgroup 3C is comprised of AR, GR, MR, PR-A and PR-B (Figure 1.1.). (Subgroup 3B is comprised of the estrogen-related receptors, which share a high degree of sequence identity with ER-\(\alpha\) and ER-\(\beta\), but do not bind steroid hormones\textsuperscript{8}. ) One major difference between the two subgroups is that subgroup 3A members recognize the consensus sequence AGGTCAnnnTGACCT and the 3C members recognize the consensus sequence AGAACAAnnnTGTTCT.

![Phylogenetic tree of the steroid receptor family](image)

**Fig. 1.1.** Phylogenetic tree of the steroid receptor family depicting the family’s divergence. The filled circle represents an ER-like common ancestor for subgroup 3A (ER-\(\alpha\) and ER-\(\beta\)) and subgroup 3C (PR, AR, GR, and MR). Because PR isoforms are generated from the same gene, the isoforms are not shown.

The receptors share a common modular structure consisting of a highly-conserved DNA-binding domain (DBD), a ligand-binding domain (LBD), and N-terminal region (NTR). Housed within the NTR and LBD are activation functions 1 and 2, respectively, which recruit co-activating proteins. The standard model of steroid receptor function is as follows. Steroids, small hydrophobic molecules, passively diffuse through the cell membrane and bind to their cognate receptors in the cytoplasm. Upon binding steroid, the steroid receptor undergoes a conformational change to release heat shock proteins. Steroid-bound receptors then dimerize,
translocate to the nucleus, and bind specific DNA sequences in the promoter and enhancer regions of a gene. The receptors then recruit co-activating proteins and the general transcriptional machinery to initiate the production of messenger RNA. The standard model provides a general framework of steroid receptor function; however, a quantitative understanding is lacking. In this thesis, I first present how my studies on the androgen receptor improve our understanding of how members of a homologous family of transcription factors achieve their specific functions. Second, I present how my studies on the glucocorticoid receptor improve our understanding of the dynamic assembly of GR at its response elements.

**How Do the Steroid Receptors Perform Their Specific Functions?**

An open question in the field is how the members of the steroid receptor family recognize the same HREs yet regulate distinct networks of genes. For example, the subgroup 3C receptors recognize the same consensus sequence, yet the receptors regulate distinct networks of genes. The androgen receptor regulates genes that contribute to the maintenance and development of the male phenotype; whereas, GR regulates genes associated with the stress response and circadian rhythms. Further, PR regulates female fertility genes and MR regulates genes associated with electrolyte and water homeostasis. Given that the subgroup 3C receptors recognize identical response elements, how are these distinct biological functions accomplished?

In general, there are two schools of thought that address this question. In the first school, interactions on a higher order than receptor-DNA interactions are responsible for steroid receptor-specific function. One class of higher-order interactions that have been demonstrated to give rise to receptor-specific function is the way by which the receptors differentially interact with chromatin structure. Eukaryotic DNA is organized and packaged into chromatin, which acts
as a general suppressor of gene activation. Chromatin structure achieves suppression by controlling the accessibility of DNA-binding proteins for their DNA-binding sites, which ultimately prevents the assembly of the pre-initiation complex. The pre-initiation complex recruits RNA polymerase and marks the initiation of transcription. To study the effects of chromatin structure on transcriptional activation by the steroid receptors, the left terminal repeat of the mouse mammary tumor virus promoter (MMTV-LTR) promoter has been used as a model system. When stably integrated into the chromatin structure of cells, the MMTV-LTR promoter is positioned in such a way that four HREs are positioned on a nucleosome (Nuc-B).

Upon induction by glucocorticoids and progestins, the DNA associated with Nuc-B becomes hypersensitive to restriction endonucleases, chemical agents, and endonucleases.

These results suggest that the steroid receptors interact directly with chromatin-bound DNA and recruit co-activating proteins with chromatin remodeling functions. To test if steroid receptors interact directly with histone-bound DNA, DNase footprint and the electrophoretic mobility shift assay (EMSA) have been used with nucleosomes reconstructed in vitro and purified DNA-binding domains of AR, PR, and GR. Results indicate that the receptors are capable of interacting directly with nucleosomal DNA.

Although the steroid receptors can directly interact with nucleosomal DNA, do the receptors differentially interact with chromatin structure once bound to elicit different responses? To address this question, the MMTV promoter fused to the chloramphenicol acetyltransferase (CAT) reporter gene was stably integrated into chromatin and transcriptional activation was measured after induction with glucocorticoids and progestins. Interestingly, glucocorticoids conferred a strong response; whereas, progestins elicited a very weak response. One reason for this difference could be that GR was able to recruit the necessary chromatin remodeling proteins for the particular chromatin structure and PR was not able to...
perform this task. This study demonstrates that chromatin structure is one way by which steroid receptors can recognize the same set of hormone response elements yet regulate distinct networks of genes.

A second class of higher-order interactions by which steroid receptors exert their specific function is through co-activating proteins. The steroid receptors have been shown to interact directly with components of the basal transcriptional machinery, such as TFIID/TBP, TFIIF and TFIIH; however, interactions with co-activating proteins have been attributed to playing a more significant role. One family of co-activating proteins in particular is the steroid receptor co-activator (SRC) family of co-activating proteins. This co-activator family includes SRC-1, SRC-2 (TIF2/GRIP-1), and SRC-3 (pCIP/RAC3/ACTR/TRAM1). The modular structure of the SRC family members consists of nuclear receptor (NR) boxes that interact directly with steroid receptors. One well characterized motif is the LXXLL motif, where L represents a lysine residue and X represents any amino acid. The LXXLL region has been demonstrated to interact directly with activation function 1 (AF-1) located in the receptor ligand-binding domain. The SRC structure also consists of a domain with histone acetyltransferase (HAT) activity. Upon acetylation, it is generally thought that the association between DNA and a histone is weakened, allowing for access of additional DNA-binding proteins that lead to transcriptional activation of a gene.

The SRC members also contain a region(s) that can interact with additional chromatin remodeling proteins. At a stably integrated MMTV promoter, DNA-bound PR interacts with SRC-1, which then goes on to recruit CREBS-binding protein (CBP). CBP then acetylates a histone and “loosens” the nucleosomal structure. On the other hand, DNA-bound GR interacts with SRC-2, which then recruits pCAF. pCAF then acetylates a different histone and “loosens” the promoter in a different location of the promoter region. Thus, co-activating proteins could serve
as factors that confer receptor-specific function. However, most co-activating proteins have been shown to interact with more than one receptor and may not directly confer receptor-specific function\textsuperscript{16}. Co-activating proteins are currently thought to be a part of a large assembly of proteins, each of which is modulated by a variety of signaling pathways\textsuperscript{20}. One can imagine that the activity levels of the large assembly is governed by the concentrations of the component proteins and by post-translational modifications of the different component proteins. Although one can imagine this assembly could influence receptor specificity, the mechanism by which the vast diversity of protein-protein interactions and post-translational modifications influence receptor specificity have not been demonstrated.

Another group of non-receptor proteins that have been shown to influence specificity are DNA-binding auxiliary proteins. At a minimal promoter with only one hormone response element, GR and MR equally activate a reporter gene. This changes, however, once a more complex promoter is introduced. The plfG gene promoter is one such example. This promoter consists of a HRE and a site that can bind an auxiliary protein, activator protein 1 (AP-1), which is a heterodimer of cJun and cFos. The ratio of the cJun and cFos has been shown to differentially influence GR and MR function. These results suggest that non-receptor DNA-binding proteins can influence receptor-specific function\textsuperscript{21}.

A second example of an auxiliary factor influencing receptor-specific function has been observed at the promoter of the mouse sex limited protein (Slp). Slp has been shown to be induced by androgens but not by glucocorticoids or progestins, but only when a 120 bp upstream region is present. This upstream region contains binding sites for non-receptor auxiliary factors. Even though GR was shown to bind the promoter, the receptor does not induce transcriptional activation of the gene. Taken together, these two examples demonstrate
that receptor-specific responses could be heavily influenced by steroid receptor-auxiliary factor
interactions that are dictated by the context of the promoter.22

The second school of thought surrounding steroid receptor-specific function posits that
lower-level interactions between the receptors and DNA govern specificity. This school claims
that interactions of receptor-DNA interactions possess the ability to confer specificity among the
receptor family members. As mentioned above, the subgroup 3C receptors (AR, GR, PR, and MR)
recognize the same consensus hormone response element (HRE) comprised of two six-base-pair
half-sites separated by a three-base-pair spacer. Recent sequencing studies demonstrate that
these receptors can bind to a broad range of HRE sequences. The function of these subtle
differences in HRE sequence as they relate to receptor specificity is currently under debate. One
hypothesis claims that the steroid receptors possess the ability to modulate their structure upon
binding a particular HRE sequence.23 The modified receptor structure, encoded by the response
element sequence, recruits a particular co-activating protein or a particular assembly of co-
activating proteins with a response element-dictated affinity. When bound to a slightly different
HRE sequence, the holoreceptor will take-on a different conformation and will display a
different affinity for its co-activating proteins and could potentially recruit a different set of co-
activating proteins. In short, DNA is an allosteric effector of steroid receptor conformation,
which dictates function.

The majority of evidence supporting the “allostERIC hypothesis” comes from studies with
the GR DBD. Crystallography and NMR studies indicate that two noticeable changes occur at the
DBD dimer interface and a distorted region between the two zinc fingers, coined the lever
arm.24,25 Subsequent studies demonstrate that a panel of response elements confer different
conformations of the lever arm. DNA-binding studies indicate that the GR DBD displays a broad
spectrum of DNA-binding affinities and suggest that affinity is not correlated with transcriptional
activation as measured with full-length GR in a reporter gene assay. One reason for sequence-specific transcriptional activation could be that the different response element sequences confer different conformations in the GR DBD, which alter the structure of the holoreceptor, and modulate the functionality of the protein. Given that GR shares less than 20% sequence identity in the N-terminal region, which makes-up a large portion of each receptor and houses activation function 1 (AF-1), one can imagine that each receptor will have its own way of interpreting the conformation code prescribed by response element sequence. Thus, response element sequence is a possible means by which the receptors could attain their specific functions.

Response element sequence has been seen to confer specificity to the androgen receptor (AR). Instead of containing the inverted palindrome sequence of the consensus sequence, the AR-regulated rat probasin (PB) and secretory component (SC) contain response elements containing a direct repeat sequence (e.g. AGAACA\textit{nnnn}AGAACA). DNA-binding studies demonstrate that the AR DBD binds with a significantly stronger affinity than the GR DBD. These studies raise the possibility that receptor-specific response elements can give rise to receptor-specific occupancy, which ultimately gives rise to receptor-specific function.

As mentioned above, previous studies with the GR DBD suggest that DNA-binding affinity at a panel of response element sequences is not correlated to transcriptional activation. Recent work in our lab, however, indicates the contrary. By examining the correlation between DNA-binding affinity and transcriptional activation with a different statistical approach, the lab found that GR DNA-binding affinity is a primary determinant of sequence-specific transcriptional activation. Given that DNA-binding affinity is a primary determinant of GR function, the interactions between the steroid receptor members and DNA may play a role in receptor specificity.
To investigate the thermodynamic mechanisms of steroid receptor-specific function, our lab is systematically dissecting the energetics of receptor-promoter interactions at a model two-site promoter. We hypothesize that differences in the microscopic energetics of promoter assembly among the receptors give rise to steroid receptor-specific function. This approach has provided one explanation for how the two isoforms of the progesterone receptor (PR-A and PR-B) achieve their specific functions. The two isoforms are identical, with the exception that PR-B contains an additional 164 amino acids at the N-terminus. Connaghan and Heneghan found that both receptors dimerize with roughly micromolar affinities. Regarding DNA-binding interactions, PR-A binds the model promoter with negligible intersite cooperativity between DNA-bound receptor dimers; whereas PR-B binds with substantial cooperativity. Relative to PR-B, this disparity in cooperativity diminishes PR-A occupancy as the promoter becomes saturated. Because increased intersite cooperativity will increase the probability of observing a fully occupied promoter, these results predict that PR-B will regulate the bulk of progesterone-responsive genes. Indeed, a microarray study (cell type) demonstrated that of 94 genes, 65 are regulated by PR-B, 4 by PR-A and 25 by both isoforms. Thus, a rigorous thermodynamic analysis of receptor-specific DNA interactions could explain receptor-specific gene regulation in the context of a living cell.

To address the question of specificity among the entire family, this approach has been extended to some of the remaining receptors. Similar to the self-assembly and DNA-binding energetics observed for the PR isoforms, another steroid receptor subgroup 3C member, glucocorticoid receptor (GR), showed no evidence of dimerization and assembled the model two-site promoter with substantial intersite cooperativity. The more distantly related subgroup 3A member ER-α, conversely, displayed strong self-assembly energetics and negligible intersite cooperativity at the two-site promoter. Taken together, these results imply that the receptors...
allot their energetics in line with their phylogeny. For example, the subgroup 3A receptor ER-α exhibits strong dimerization and weak cooperativity; whereas the subgroup 3C receptors PR-B and GR exhibit weak (or no) dimerization and strong cooperativity. Of the remaining 3C receptors, I hypothesized that under identical buffer conditions the androgen receptor (AR) would display the same energetics as the other subgroup 3C receptors: weak dimerization and strong cooperativity.

To this end, in Chapter II I present a thermodynamic dissection at the model two-site promoter of wild type AR and a point mutation that is associated with prostate cancer progression, T877A. This point mutation is located in the steroid binding pocket of the LBD and results in an expanded repertoire of activating ligands\textsuperscript{37,38}. Carrying out this work extends our understanding of how the steroid receptor family members interact with their promoters; interactions that are fundamental to transcriptional activation events by this family. Additionally, this research allows us to examine if lower-level interactions between receptor and DNA can be used to help explain transcriptional activation.

Under a specified or established set of experimental conditions (e.g. buffer composition and temperature) we are dissecting the interactions of the steroid receptor family members at a two-site model promoter. This two-site promoter, denoted as HRE\textsubscript{2}, is comprised of two identical response elements from the tyrosine aminotransferase (TAT) gene. A schematic of the dimer binding pathway at this two-site promoter is depicted in Figure 1.2.a. Gibbs free energies are expressed as affinity constants for comparative purposes with the steroid receptor literature. Following the standard model of steroid receptor function, receptor monomers dimerize through equilibrium constant $k_{\text{dim}}$. A preformed dimer then binds to a response element with an intrinsic binding affinity, $k_{\text{int}}$. Both of these microscopic equilibrium constants possess a straightforward molecular interpretation. Alternatively, the binding of two monomers
to a single response element can be expressed with the macroscopic equilibrium constant $K_{\text{tot}}$. It should be noted that these values do not represent the commonly reported apparent binding affinity ($K_{\text{app}}$) resolved from the Langmuir model. Resolution of $K_{\text{app}}$ can be useful to compare the relative affinities for a panel of response elements, but $K_{\text{app}}$ is a composite of microscopic equilibrium constants (as discussed above) and does not possess a molecular interpretation. Finally, as shown in Figure 1.2.b, the binding of one receptor dimer can influence the binding of a second receptor dimer through cooperative interactions ($k_c$).

**Figure 1.2.** Representative schematic of a dimer binding pathway at the HRE2 promoter. (a) Representative dimer binding pathway of HRE2 assembly. Dimers form via the microscopic equilibrium constant $k_{\text{dim}}$. Preformed dimers then bind a single response element intrinsic affinity $k_{\text{int}}$. Regardless of pathway, the total reaction of two monomers assembling at a response element can be captured with equilibrium constant $K_{\text{tot}}$. (b) Potential intersite cooperativity between bound dimers is represented with $k_c$.

Because I hypothesized that differences in the microscopic interactions are determinants of receptor-specific function, resolving these microscopic interactions is of utmost importance to our studies. To resolve the dimerization constant ($k_{\text{dim}}$) I used analytical ultracentrifugation methods (AUC) (sedimentation velocity and equilibrium)\textsuperscript{39,40}. Additionally,
sedimentation velocity experiments were used to examine the polydispersity of steroid receptor preparations to help ensure a quality protein preparation.

To resolve the equilibrium constants depicted in Fig. 1.2, I used quantitative DNase footprint titrations and construct statistical thermodynamic models with physically meaningful parameters to fit the data. Compared to traditional methods of measuring receptor-DNA interactions (e.g. filter binding), quantitative footprinting has the added benefit of monitoring receptor binding at the individual response elements or binding sites. Without the ability to distinguish between the different binding sites, it is not possible to reliably resolve the intersite cooperativity constant $k_c$. Although the filter binding technique, electrophoretic mobility shift assay (EMSA or gel shift), and surface plasmon resonance are more common techniques, they lack the ability to distinguish interactions at specific sites for a multi-site promoter and thus cannot reliably report on the microscopic interactions at complex promoters.

Prior to my work on the two androgen receptor constructs, other groups have examined the self-assembly and DNA-binding energetics of this receptor. Relative to the other steroid receptors, AR dimerization has not been examined extensively. The majority of work on steroid receptor dimerization has focused on PR, ER and GR. This early work carried-out on PR, ER, and GR indicate that the steroid receptors exist as preformed dimers prior to binding DNA. Because the receptors share a common modular structure, it appears that AR dimerization was predicted to behave like the others and has not been studied extensively. One thing that came out of the work on the other receptors is that the ligand-binding domain contains the primary dimerization interface. Thus, the small amount of work on AR dimerization that does exist focuses on the AR LBD.

Using size exclusion chromatography, Ota and co-workers compared the sizes of GST-tagged DBD and LBD. The group found that GST-DBD formed dimers; whereas GST-LBD formed a
high molecular weight species (>400 kDa) that eluted in the void volume. The authors attribute the GST-DBD dimers to GST self-assembly, a correct conclusion. The authors then go on to claim that the high molecular mass species of the GST-LBD is due to GST-GST and LBD-LBD interactions. The authors conclude that the high molecular species provide evidence that AR forms dimers in solution (before it binds DNA). The receptors are prone to aggregation; however, the authors do not take this into consideration and may not have known at the time the research was performed.

Isolating the full-length receptors is a tall task and only one paper has examined the dimerization of full-length AR. Wilson and co-workers expressed and purified full-length human AR with the baculovirus expression system and examined the influence of redox state on AR dimerization. Using PAGE methods, the group found that AR forms disulfide-linked dimers under certain redox conditions. Although the formation of dimers required micromolar amounts of AR, the group claimed to have observed AR dimerization in bulk. These results, however, are most likely an artifact of the protein purification process. During the protein prep, the group freeze-thawed the protein preparation 3 times prior to an experiment. Multiple freeze-thaw cycles can be detrimental to a protein, especially a full-length transcription factor, and the high molecular mass protein observed by PAGE could simply be due to this shortcoming of the study.

Miesfeld and co-workers took a thermodynamic approach to examining the influence of DNA-binding affinity on AR-specific promoter occupancy. Using a reporter gene assay, the group found that AR and GR differentially activate isolated response elements from the promoters of mouse mammary tumor virus (MMTV), tyrosine aminotransferase (TAT), prostatein (C3), and sex limited protein (SLP). In an attempt to account for these differences, Misteli measured the apparent binding affinities of AR and GR DBDs with quantitative footprinting. Apparent affinities for AR ranged from 43 nM to 460 nM, and differences in apparent affinities could not account
for differences in transcriptional activation. This finding is expected since the receptors share approximately 80% sequence identity. Altogether, the existing literature indicates that AR binds DNA with an apparent affinity on nanomolar scale.

In the literature, there are no quantitative studies that examine the intersite cooperativity of full-length human AR. Previous work provides qualitative examinations of intersite cooperativity with AR DNA-binding domain constructs. Matusik and co-workers found that two response elements in the promoter region of the rat probasin (PB) gene were required for full induction, suggesting cooperative interactions between two DNA-bound AR dimers53. Using DNase footprinting, the group found that AR cooperatively binds to the two response elements in the PB promoter. Similarly, Robins and co-workers found that AR cooperatively binds to two response elements in the promoter region of the mouse sex limited protein (Slp) promoter54. Finally, subsequent work by Carey and co-workers suggested that AR cooperatively binds multiple response elements within the far upstream enhancer region of the PSA promoter55. Taken together, although qualitative in nature, the literature is in agreement; AR is capable of intersite cooperative interactions.

**What Controls the Dynamics of GR-Promoter Assembly?**

In addition to the question of steroid receptor specificity discussed above, another open question in the field is how the dynamics of steroid receptor-promoter interactions influence transcriptional activation. Transcriptional activation is a dynamic process. Dynamic interactions between transcription factors and their DNA-binding sites, co-activating proteins and the basal transcriptional machinery play roles leading to successful activation of a gene. Because the process is inherently dynamic, understanding the dynamics of the individual interactions will lead to a better understanding of the underlying mechanisms.
For the past several decades, the glucocorticoid receptor (GR) has served as a model for receptor-DNA dynamics both in the context of a living cell and with DNA alone in the test tube. Microscopy experiments with living cells demonstrate that GR-DNA interactions display residence times on the timescale of seconds. The majority of evidence comes from examinations of GR dynamics at a contiguous array of approximately 200 copies of the mouse mammary tumor virus (MMTV) promoter that was stably integrated into the chromatin structure\textsuperscript{56,57}. Although an artificial promoter architecture, this large array of the MMTV promoter was constructed so that GR dynamics could be monitored with the microscopes available at the time. Recent single molecule tracking (SMT) studies at individual natural promoters confirm the results of the MMTV array studies; GR dynamics occur on the timescale of seconds\textsuperscript{58,59}. These results indicate that, relative to the time involved in the initiation of transcription, GR-DNA dynamics are transient. In fact, it appears that GR cycles on and off its promoters in a ‘hit-and-run’ manner as the pre-initiation complex is being assembled to recruit the RNA polymerase.

These recent live-cell imaging studies came as a surprise. Prior to this work, it was generally thought that the steps leading to transcriptional activation were static. GR was previously thought to bind to a promoter, recruit co-activating proteins, and recruit the basal transcriptional machinery, all while GR is statically bound to its promoter. In part, this static model of transcriptional activation was supported by earlier work performed with the purified GR DNA-binding domain (DBD) and DNA. The first \textit{in vitro} investigation (and most heavily cited) examined GR dissociation kinetics from an isolated response element of the mouse mammary tumor virus (MMTV) promoter, the same promoter used in the live-cell imaging studies. Using time-resolved DNase footprinting, Wrange and co-workers found that GR dissociates under a one-phase exponential decay model with an observable rate constant of $1.1 \times 10^{-4}$ s\textsuperscript{-1}, which translates to a calculated residence time of 150 minutes\textsuperscript{60}. (The calculated residence time is the
inverse of the observable rate constant.) This investigation was a minor portion of an ambitious study that, in addition to dissociation kinetics, surveyed the stoichiometry, DNA-binding affinity, and intersite cooperative interactions of the GR DBD at the long-terminal repeat of the MMTV promoter. Despite this innocent shortcoming, the dissociation work has been cited extensively in the literature to claim that dissociation kinetics measured in vitro occur on the timescale of minutes to hours\textsuperscript{57,58}.

Under comparable experimental conditions, subsequent work by Lieberman and Nordeen confirmed that GR dissociates from an isolated response element on the timescale of minutes to hours. Using an electrophoretic mobility shift assay (EMSA), they found that GR dissociates under a one-phase decay model with an observable rate constant of $2.7 \times 10^{-4}$ s$^{-1}$\textsuperscript{61}. This rate constant translates to a calculated residence time of 62 min. The roughly two-fold decrease in residence time observed by Nordeen could be due to the use of the EMSA, which may underestimate the residence time. To prevent unbound GR from re-associating back onto radiolabeled DNA, both groups used unlabeled competitor DNA to sequester unbound GR. Nordeen found that as the concentration of competitor DNA was increased 200-fold, the observable rate constant increased to $3.6 \times 10^{-3}$ s$^{-1}$. This rate constant translates to a residence time of 4 minutes. As the amount of unlabeled DNA is increased, the rate of dissociation starts to approach the timescale observed in live-cell imaging studies. These results were interpreted to mean that GR undergoes intersegmental transfer between separate DNA strands. This is one mechanism of facilitated diffusion so that GR can find its binding sites in an expedient manner.

On the other hand, using surface plasmon resonance (SPR), Yamamoto and co-workers found that the GR DBD dissociates from a panel of response elements on the timescale of seconds to tens of seconds\textsuperscript{25}, the same timescale observed in living cells\textsuperscript{57,58}. SPR does not use unlabeled competitor DNA to sequester unbound GR. Instead, the flow of the experimental
buffer washes away any free GR from re-binding to the surface-bound DNA fragment. Unlike the work by Noreen, this study suggests that high concentrations of unlabeled DNA are not required for GR to dissociate on the timescale of seconds to tens of seconds. GR interacts with a variety of response elements that differ by 1 or 2 base pairs to elicit its response. Analogous to the work done by Wrange, the SPR experiments were a minor portion of work that examined the influence of response element sequence on GR structure. As such, the SPR work was not emphasized and the analysis was not fully disclosed. For example, the association kinetics are not discussed and it is unclear if the investigators observed one-phase or two-phase dissociation kinetics. Although Yamamoto examined the kinetics at a panel of response elements, a systematic approach to the dissociation kinetics of full-length GR is still lacking.

Taken together, the literature examining in vitro dissociation is contradictory. For example, studies performed with the GR DBD by Wrange and Nordeen indicate that GR dissociates on the timescale of minutes to hours, which is in stark contrast to the live-cell imaging studies. This discrepancy has been interpreted to mean that GR dissociation on the second timescale requires processes unique to an intact cell. On the other hand, work by Yamamoto indicates that GR dissociates in seconds to tens of seconds, the same timescale observed in living cells. The purity and activity of the receptor subdomain are unclear for some of these studies, which could influence the contradictory in vitro results. Additionally, it should be noted that the isolated DBD, which makes-up roughly 10% of the protein, was used for the in vitro studies; whereas live-cell imaging studies use the full-length holoreceptor. These observations call for a re-examination of the in vitro GR-DNA dissociation kinetics. Doing so will accomplish several things. First, it will help reveal if the receptor can dissociate quickly on its own or if it requires additional processes or factors to do so. Second, a rigorous in vitro kinetic analysis also allows for the examination of the mechanisms by which GR interacts with its DNA-
binding sites, something live-cell imaging studies are not yet able to achieve. To this end, I performed a systematic analysis of GR dissociation kinetics with full-length human GR that was shown to be amenable to a rigorous analysis. Experiments were performed over a broad range of protein concentrations, used six response element sequences and a multisite promoter, and used a chimeric GR construct. Mechanisms elucidated with \textit{in vitro} techniques may not necessarily match the mechanisms used by an intact cell, but provide models in which we can place under consideration.

\textbf{Overview}

This thesis is divided into four chapters, including this introduction, and an appendix. In Chapter II, I will present my work on the thermodynamics of promoter assembly with two androgen receptor constructs as a way to demonstrate my contribution to our understanding of steroid receptor-specific function. In Chapter III, to establish my contribution to our understanding of transcription factor-promoter dynamics, I will present my work on the \textit{in vitro} dissociation kinetics of the glucocorticoid receptor from a panel of six hormone response elements and the natural mouse mammary tumor virus (MMTV) promoter. In Chapter IV, I will summarize the work performed in the middle two chapters, frame how these studies fit within the larger context of the field, and provide recommended future directions for continuing these projects. In the appendix, I will address two possible shortcomings associated with the \textit{in vitro} studies of GR-DNA dissociation kinetics.
Summary

Steroid receptors comprise a homologous family of ligand-activated transcription factors. The members include androgen receptor (AR), estrogen receptor (ER), glucocorticoid receptor (GR), mineralocorticoid receptor (MR) and progesterone receptor (PR). Phylogenetic studies demonstrate that AR, GR, MR and PR are most closely related, falling into subgroup 3C. ER is more distantly related, falling into subgroup 3A. To determine the quantitative basis by which receptors generate their unique transcriptional responses, we are systematically dissecting the promoter-binding energetics of all receptors under a single “standard state” condition. Here we examine the self-assembly and promoter-binding energetics of full-length AR and a mutant associated with prostate cancer, T877A. We first demonstrate that both proteins exist only as monomers, showing no evidence of dimerization. Although this result contradicts the traditional understanding that steroid receptors dimerize in the absence of DNA, it is fully consistent with our previous work demonstrating that GR and two PR isoforms either do not dimerize or dimerize only weakly. Moreover, both AR proteins exhibit substantial cooperativity between binding sites, again as seen for GR and PR. In sharp contrast, the more distantly related ER-α dimerizes so strongly that energetics can only be measured indirectly, yet cooperativity is negligible. Thus homologous receptors partition their promoter-binding energetics quite differently. Moreover, since receptors most closely related by phylogeny partition their

---

1 Chapter II reproduced with permission from De Angelis RW, Yang Q, Miura MT, Bain DL. *J Mol Biol.* 2013;425:4223-4235. Copyright 2013 Elsevier Ltd.
energetics similarly, such partitioning appears to be evolutionarily conserved. We speculate that such differences in energetics, coupled with different promoter architectures, serve as the basis for generating receptor-specific promoter occupancy and thus function.

Introduction

The androgen receptor (AR) is a member of the steroid receptor family of ligand-activated transcription factors\(^3\). The remaining members include the estrogen receptor (ER), glucocorticoid receptor (GR), mineralocorticoid receptor (MR) and the progesterone receptor (PR). ER exists naturally as two isoforms (ER-\(\alpha\) and ER-\(\beta\)) as does PR (PR-A and PR-B).

Phylogenetic studies indicate that although all receptors descend from a common ER-like ancestor, the family eventually branched into two distinct subgroups\(^7\). As shown in Figure 2.1.a, these correspond to subgroup 3A (ER-\(\alpha\) and ER-\(\beta\)) and 3C (AR, GR, MR and PR)\(^8\). One of the key differences between the subgroups is that the ER isoforms recognize a consensus AGGTCAnnnTGACCT hormone response element (HRE), whereas the more closely related AR, GR, MR and PR recognize a consensus AGAACAnnnTGTTCT.

The generally accepted model of receptor function is that upon binding hormone, the receptors dimerize via their C-terminal hormone-binding domains, bind HREs within upstream promoter sites, and recruit coactivator proteins to activate transcription\(^3\). Although this model provides a strong qualitative understanding of receptor function, it nonetheless remains incomplete. Specifically lacking is a quantitative framework for considering receptor-specific gene regulation: How does a homologous family of transcription factors, capable of binding identical or nearly identical response elements, regulate different gene networks? Although we have qualitative insight into aspects of this question, a quantitative and mechanistic understanding has yet to be determined.
Figure 2.1. Steroid receptor phylogenetic tree and promoter binding energetics. (a) Phylogenetic tree representing divergence of the steroid receptor family. Filled circle represents the ER-like common ancestor for subfamily 3A (ER-α and ER-β) and subfamily 3C (PR, AR, GR and MR). The two PR isoforms are not shown since they are generated from the same gene via alternative transcriptional or translational start sites. (b) Representative dimer-binding pathway for receptor-HRE\textsubscript{2} assembly. Dimers are formed through microscopic equilibrium constant \(k_{\text{dim}}\) and pre-formed dimers bind response elements with intrinsic affinity \(k_{\text{int}}\). Schematic representing the overall reaction for receptor monomer assembly at a palindromic response element regardless of pathway. Total affinity is represented by \(K_{\text{tot}}\). (c) Potential inter-site cooperativity is represented by \(k_{c}\).

As a step toward addressing receptor-specific transcriptional behavior, we are systemically dissecting the promoter-binding energetics of all the human steroid receptors. We are attempting to do so at a single “standard state” condition under which the receptors are amenable to rigorous analysis and thus meaningful comparison (pH 8.0, 100 mM NaCl, 4 °C). Our rationale is based on previous work demonstrating that receptor-DNA interaction energetics are the primary determinant of sequence-specific transcriptional activity within the
This finding suggests that energetics should play an important role in generating receptor-specific function as well. However, only by dissecting out each individual energetic contribution to receptor-promoter assembly (receptor dimerization, DNA binding, cooperative stabilization and coactivator recruitment, for example) is it possible to identify which parameter(s) might be most critical to specificity. By comparison, more traditionally measured terms such as an apparent affinity ($K_{app}$) or Hill Coefficient ($n$) are composite values typically lacking in molecular insight.

Shown in Figure 2.1.b are representative assembly states and interaction energetics for receptor assembly at a simple promoter containing two hormone response elements (HRE$_2$). Energetics are expressed as affinity constants to facilitate comparison to other reports. By the traditional model, receptors dimerize in the absence of DNA ($k_{dim}$) and bind to response elements as pre-formed dimers ($k_{int}$). These terms – all microscopic interaction constants with clear molecular interpretations – can also be expressed as part of a macroscopic constant, $K_{tot}$. This parameter describes the total affinity of loading two monomers onto a response element, regardless of whether assembly occurs via a pre-formed dimer as described here or via successive assembly of monomers$^{63}$. Thus using the values in Figure 2.1.b, $K_{tot} = k_{dim} \cdot k_{int}$. Finally, as shown in Figure 2.1.c, assembly at a multi-site promoter such as HRE$_2$ may also be coupled to inter-site cooperativity ($k_c$).

In our earlier studies of steroid receptor assembly at the HRE$_2$ promoter, we first noticed that dimerization energetics ($k_{dim}$) vary enormously. For example, if expressed as dissociation constants, ER-α dimerizes with an affinity of 0.35 nM under conditions in which the PR isoforms dimerize with values of 1–2 μM, or ~ 1,000-fold weaker$^{31,34,35}$. Surprisingly, GR
showed no evidence for dimerization, allowing us to place only a lower limit of 100 μM, or at least 100,000-fold weaker than ER-α\textsuperscript{32}. Inter-site cooperative binding ($k_c$) also varied significantly and trended inversely with dimerization. For example, ER-α exhibits essentially no cooperativity ($k_c = 1.4$) whereas GR and PR-B maintain substantial cooperative stabilization ($k_c = 70$)\textsuperscript{31,32}. Thus, ER-α, a receptor affiliated with subgroup 3A, partitions its interaction energetics in such a way to generate exceptionally strong dimerization but little or no cooperativity. By contrast, receptors in subgroup 3C such as GR and PR-B exhibit distinct behavior – dimerization is weak or non-existent but inter-site cooperativity is strong.

These findings demonstrated that: 1) a homologous family of receptors partitions its promoter-binding energetics quite differently, 2) differences in partitioning span many orders of magnitude, and 3) receptors most closely related in phylogeny partition their energetics similarly. Taken together, these results suggested to us that differences in promoter binding energetics are an evolutionarily conserved feature of the steroid receptor family, and thus critical to receptor-specific function. Such an interpretation further predicts that human AR, a subgroup 3C receptor, should display promoter-binding energetics most similar to those of other 3C receptors and distinct from ER-α. To test this, we rigorously dissected the self-assembly and promoter-binding energetics of full-length human AR under conditions identical to those of our earlier studies. To independently confirm our results, we also analyzed T877A, a well-known functional mutant associated with advanced prostate cancer\textsuperscript{37,38}. T877A is one of the few mutations found in prostate cancer for which the functional change in mechanism is well understood, and the protein itself is amenable to biochemical and structural analysis\textsuperscript{64}.
Using analytical ultracentrifugation, we find that both AR proteins exist only as monomers, showing no evidence of dimerization up to and above micromolar concentrations. Using quantitative footprint titrations, we find that both proteins also exhibit strong inter-site cooperativity. Thus as predicted, AR partitions its promoter-binding energetics in line with other closely related family members. This finding suggests a basis for driving receptor-specific transcriptional control. Specifically, large differences in receptor-specific energetics define a series of receptor-specific energetic profiles; coupling these profiles to different promoter architectures may create a molecular framework for generating receptor-specific promoter occupancy and thus receptor-specific gene regulation. Importantly, such a framework predicts preferential promoter binding even if multiple receptors are competing for identical response elements, as is the case for the 3C subgroup.

Results

Wild-type (wt) AR and T877A were expressed as hexahistidine-tagged proteins in baculovirus-infected Sf9 cells. Each receptor was isolated to greater than 90 % purity as judged by quantification of Coomassie-stained SDS-PAGE (Figure 2.2.a). Average yields were 1 mg/L of cell culture. Since we found that even a single freeze-thaw cycle initiated irreversible AR aggregation and polydispersity, only freshly prepared receptor was used for all analyses. Finally, although we have not directly tested whether the hexa-histidine tag influences AR function, our work on ER-α, GR, and both PR isoforms has shown that neither a FLAG-tag nor His-tag has any effect\textsuperscript{29–32,35,65,66}. 


Figure 2.2. Purification of full-length human AR and characterization using sedimentation velocity and native-PAGE. (a) Baculovirus-expressed wt AR and T877A were purified as described in Materials and Methods. Purified receptors were resolved using 10 % SDS-PAGE and Coomassie-Blue staining. Molecular mass markers are indicated on left. (b) Sedimentation velocity data for 1.4 µM wt AR in 500 mM NaCl. Open circles represent absorbance data collected at 230 nm as a function of time and radial distance. Solid lines represent best fit from c(s) analysis as implemented in the program Sedfit. For clarity, only every seventh data point is displayed. (c) c(s) distributions for three concentrations of wt AR and T877A in 500 mM NaCl. Thick solid line (1.4 µM), dashed line (0.7 µM) and thin solid line (0.14 µM). (d) Immunoblot of wt AR using native-PAGE. Lower molecular mass species is indicated by large arrow; larger mass species indicated by small arrow.
Analytical ultracentrifugation demonstrates that AR exists as a monomer

Sedimentation velocity was first used to characterize the hydrodynamic and self-association properties of wt AR and T877A. Like all other receptors we have examined, AR is most soluble at high salt concentrations. Therefore, we first sedimented both proteins at 500 mM NaCl and over a ten-fold protein concentration range. Shown in Figure 2.2.b are representative absorbance scans of wt AR at 1.4 μM; the solid lines represent the best fit by c(s) analysis. Figure 2.2.c shows the corresponding sedimentation coefficient distributions determined for wt AR and T877A at concentrations ranging from 0.14 to 1.4 μM. It is evident that both receptors sediment as a single species regardless of protein concentration, with a buffer and temperature-corrected sedimentation coefficient ($s_{20,w}$) of 4.8 S. The absence of a concentration-dependent change in the sedimentation coefficient indicates neither receptor undergoes reversible self-association over this concentration range. With regard to assembly-state, c(M) analysis of the 4.8 S peak yields an average molecular mass of 94 ± 5 kDa for wt AR and 84 ± 14 kDa for T877A, indicative of AR monomer (calculated molecular mass of 104 kDa).

For both proteins, we also observe two minor peaks at ~2.8 and 6.5 S. Each peak represents ~5% of the total c(s) distribution regardless of receptor concentration, indicating that they represent irreversibly formed conformers or aggregate. The 2.8 S species undoubtedly represents a misfolded or partially unfolded monomer, since simple hydrodynamic calculations show that a fully unfolded, random coil AR dimer could sediment no slower than 3.4 S. Using a similar approach, we find that the 6.5 S species must represent a higher order aggregate since it sediments too quickly to represent an AR monomer, regardless of conformation. To test this latter interpretation biochemically, we electrophoresed wt AR under non-denaturing conditions (Figure 2.2.d). As seen by immunoblot analysis, we do indeed detect a high molecular mass species (small arrow) in addition to the primary AR species (large arrow). Quantification of the
former species indicates that it represents 15% of the population, in reasonable agreement with the percentage seen by sedimentation velocity. We address the exact stoichiometry of this larger species in the sedimentation equilibrium studies described below.

To rigorously determine the molecular mass and assembly-state of wt AR and T877A, we carried out sedimentation equilibrium studies under conditions identical to those of sedimentation velocity. Shown in Figure 2.3.a are sedimentation equilibrium scans for wt AR carried out at three concentrations and three rotor speeds. Analogous scans for T877A are shown in Figure 2.3.b. For both receptors, fitting the data either individually or globally to a single species model always resolved a molecular mass slightly above AR monomer. However, the quality of the fit was poor and generated non-random residuals (data not shown). Global fitting using equilibrium self-association models (monomer-dimer, monomer-trimer, etc) resolved similarly poor fits and non-random residuals.

Based on the sedimentation velocity results indicating the presence of an irreversible 6.5 S aggregate, we therefore fit the data to a non-interacting, two species model. (It is unnecessary to account for the 2.8 S partially unfolded monomer, since being the same mass as the 4.8 S monomer it would be invisible by sedimentation equilibrium). As seen by the lines through the data and random residuals, this model readily described all datasets for both receptors. Moreover, for wt AR, the fit resolved a species 1 molecular mass of 104 ± 1 kDa indicative of AR monomer, and a stoichiometry for species 2 of 3.0 ± 0.1, indicative of AR trimer. An essentially identical result was seen for T877A, resolving a species 1 molecular mass of 104 ± 1 kDa and a stoichiometry of 2.6 ± 0.2. These results were invariant of analysis regardless of how the data were edited or which subsets of data were analyzed. Thus sedimentation velocity and sedimentation equilibrium studies independently confirm that wt AR and T877A exist only as
monomers with a small amount of irreversibly formed trimer – contrary to dogma we find no evidence of AR dimerization.

**Figure 2.3.** Sedimentation equilibrium analysis of wt AR and T877A in 500 mM NaCl. (a) wt AR sedimentation equilibrium data plotted as absorbance versus $r^2/2$ for three different concentrations of AR. From left to right: 0.6 µM, 0.4 µM, and 0.3 µM. Symbols represent three different rotor speeds: 14,000 rpm (open circles), 17,000 rpm (inverted triangles), and 21,000 rpm (open squares). Solid lines represent best global fit to a two species non-interacting model. Standard deviation of global fit was 0.0039 absorbance units. Residuals are plotted below the data and best-fit lines. For clarity, only every third data point is displayed. (b) T877A sedimentation equilibrium data plotted as absorbance versus $r^2/2$ for three different concentrations. From left to right: 1.3 µM, 0.5 µM, and 0.3 µM. Symbols, corresponding rotor speeds and solid lines are identical to those described above. The standard deviation of global fit was 0.0043 absorbance units. Residuals are plotted below. For clarity, only every third data point is displayed.
**Low salt sedimentation studies confirm the presence of AR monomer and absence of dimer**

To determine AR assembly stoichiometry under our “standard state” low salt conditions, we repeated the sedimentation velocity studies at 100 mM NaCl. Shown in Figure 2.4 are the c(s) results for wt AR and T877A. For both proteins, the major species now sediments with an $s_{20,w}$ of 5.1 S (wt AR) to 5.0 S (T877A) regardless of receptor concentration. Thus the decrease in NaCl concentration causes AR to undergo slight compaction. It is also evident that the change in buffer conditions also causes the two minor peaks to increase from 5 to 10% of the total c(s) distribution. Their increased presence increases the overall heterogeneity of the AR sample, making c(M) analysis less reliable. However, assuming the 5.0-5.1 S species is indeed the AR monomer, the frictional coefficient for wt AR is calculated to be $9.5 \times 10^{-8}$ g/s and $9.6 \times 10^{-8}$ g/s for T877A. The Stokes radii are 50 Å and 51 Å for wt and T877A, respectively. Modeling the 5.0-5.1 species as a monomeric prolate ellipsoid predicts an axial ratio of 8:1. The high level of asymmetry seen for both proteins has been observed previously for ER-α, GR and both PR isoforms; it is most consistent with biochemical and structural studies demonstrating that the N-terminal regions of these proteins are natively unfolded. Finally, we speculate that the decreased NaCl concentration reduces ionic screening within and between AR monomers, thus promoting aggregation and formation of the secondary species.

To confirm our interpretation that AR is again only monomeric at 100 mM NaCl, we attempted sedimentation equilibrium studies. Under these low salt conditions, wild-type AR slowly and irreversibly aggregates over the multi-day time frame required for sedimentation equilibrium, making it impossible to collect reliable data. We were however able to collect complete data sets for T877A. We again globally fit the data using a non-interacting, two-species model. All data were well described using such an approach, resolving a species 1 molecular mass of $96 \pm 2$ kDa and a stoichiometry of $5.0 \pm 0.3$ (Figure 2.5). However, the data collected at
100 mM NaCl were not as robust as the 500 mM data, thus the stoichiometry of species 2 ranged from three to seven (trimer to heptamer), depending upon how data were edited or which subsets were analyzed. However, there was never a change in the species 1 molecular mass, indicating that under our standard state conditions, T877A exists as a monomer with a slightly increased amount (~10%) of an irreversible, non-dimeric aggregate. Noting that wt AR displayed an essentially identical c(s) distribution, we are confident in stating that it too is a monomer with a similar amount and type of aggregate. A summary of all hydrodynamic parameters for wt AR and T877A is shown in Table 2.1.

Figure 2.4. Sedimentation velocity analysis of wt AR and T877A in 100 mM NaCl. (a) c(s) distributions for three concentrations of wt AR (top). Thick solid line (1.4 µM), dashed line (0.7 µM) and thin solid line (0.14 µM). (b) c(s) distributions for three concentrations of T877A. Thick solid line (1.1 µM), dashed line (0.7 µM) and thin solid line (0.11 µM).
Figure 2.5. Sedimentation equilibrium analysis of T877A in 100 mM NaCl. T877A sedimentation equilibrium data plotted as absorbance versus \( r^2 / 2 \) for three different concentrations. From left to right: 1.3 µM, 0.5 µM, and 0.3 µM. Symbols represent three different rotor speeds: 14,000 rpm (open circles), 17,000 rpm (inverted triangles), and 21,000 rpm (open squares). Solid lines represent best global fit to a two species non-interacting model. The standard deviation of the fit was 0.0032 absorbance units. Residuals plotted as the change in absorbance versus \( r^2 / 2 \) are plotted below. For clarity, only every third data point is displayed.

Table 2.1. Wild-type AR and T877A hydrodynamic parameters

<table>
<thead>
<tr>
<th></th>
<th>500 mM NaCl</th>
<th>100 mM NaCl</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WT</td>
<td>T877A</td>
</tr>
<tr>
<td>( s_{20,w} )</td>
<td>4.8</td>
<td>4.8</td>
</tr>
<tr>
<td>( f ) (g/s)</td>
<td>( 1.0 \times 10^{-7} )</td>
<td>( 1.0 \times 10^{-7} )</td>
</tr>
<tr>
<td>( f/f_0 )</td>
<td>1.53</td>
<td>1.51</td>
</tr>
<tr>
<td>Stokes radius (Å)</td>
<td>53</td>
<td>53</td>
</tr>
<tr>
<td>axial ratio</td>
<td>10:1</td>
<td>10:1</td>
</tr>
<tr>
<td>molecular mass (kDa)(^a)</td>
<td>94 ± 5</td>
<td>84 ± 14</td>
</tr>
<tr>
<td>molecular mass (Da)(^b)</td>
<td>103,629 ± 1,846</td>
<td>104,041 ± 1,657</td>
</tr>
</tbody>
</table>

\(^a\)Average and standard deviation estimated by \( c(M) \) analysis of the major peak as implemented in Sedfit.

\(^b\)Resolved molecular mass from sedimentation equilibrium.
**AR assembles onto a two-site promoter with substantial cooperativity**

We next investigated the energetics of wt AR and T877A binding to the HRE₂ promoter using quantitative footprint titrations. Shown in Figure 2.6.a is a representative titration of the HRE₂ promoter with wt AR. Receptor binding is highly specific, and sequencing studies confirm that the protected regions correspond to the two HREs. Receptor-induced hypersensitivity adjacent to each site is also apparent. Hypersensitive regions have been observed in our work on the other steroid receptors and are likely due to receptor-induced DNA bending\(^{70,71}\).

The individual-site binding isotherms for wt AR and T877A assembly at the HRE₂ and HRE₁ promoters are shown in Figure 2.6.b. Our standard approach to quantifying receptor interactions at these promoters has been to resolve the energetics of preformed dimer assembly at each palindromic binding site: As shown schematically in Figure 2.1.b and described in more detail elsewhere\(^{29}\), a dimerization constant (\(k_{\text{dim}}\)) is independently determined by analytical ultracentrifugation and then used as a fixed parameter for the footprint titration data, allowing resolution of an intrinsic binding affinity (\(k_{\text{int}}\)). Unfortunately, this approach was not possible here since we observed no evidence of AR dimerization. Nor was a model in which AR monomers sequentially bind to palindromes, since monomer binding to an individual palindrome is highly cooperative. (Fitting the AR:HRE₁ data to the Hill Equation resolved a Hill coefficient of 2.0 ± 0.2 (data not shown)). Instead, we globally fit the HRE₂ and HRE₁ binding isotherms to Eq 2.5 and 2.6, allowing us to resolve the total affinity for assembling two AR monomers on a single palindromic HRE regardless of pathway (\(K_{\text{tot}}\)), and the cooperativity between AR dimers bound to adjacent HREs (\(k_c\)).
Figure 2.6. Quantitative DNase footprint titrations and individual site-binding isotherms of wt AR and T877A assembly at the HRE$_2$ promoter. (a) Representative autoradiogram of wt AR binding at the HRE$_2$ promoter in 100 mM NaCl. AR concentration increases from left to right. Positions of site 1 (solid rectangle) and site 2 (open rectangle) are depicted to right. (b) Individual site-binding isotherms constructed from analysis of wt AR and T877A footprint titration images. Filled red squares represent binding to site 1 and open red squares represent binding to site 2 of the HRE$_2$ promoter. Open blue circles represent binding to site 2 of the HRE$_1$ promoter. Red and blue lines represent best global fits to all isotherms using Eqs (2.5) and (2.6). The fit lines for each HRE$_2$ site overlay since the sequences of sites 1 and 2 are identical.
For wt AR, we found that the receptor assembles onto a palindromic HRE with a $K_{\text{tot}}$ of $2.1 \pm 0.2 \times 10^{13} \text{M}^{-2}$. This corresponds to a binding free energy of $-16.9 \pm 0.1 \text{ kcal/mol}$. By contrast, the apparent affinity of binding ($K_{\text{app}}$) is 220 nM or $-8.4 \text{ kcal/mol}$. The reason for this large discrepancy is that $K_{\text{app}}$ does not take into account the dimeric stoichiometry of AR assembly at the palindromic HRE. T877A generated comparable results, with a $K_{\text{tot}}$ of $9.0 \pm 0.2 \times 10^{12} \text{M}^{-2}$ or $-16.4 \pm 0.1 \text{ kcal/mol}$. Finally, noting the leftward shift and increased steepness of the HRE$_2$ isotherms relative to HRE$_1$, there must be significant cooperativity between binding sites. Global fitting of the wt AR data resolves a cooperativity term ($k_c$) of $69 \pm 15$, which translates to a free energy of $-2.3 \pm 0.1 \text{ kcal/mol}$. For T877A, the value was $96 \pm 35$ or $-2.5 \pm 0.2 \text{ kcal/mol}$. The energetic parameters for wt AR and T877A promoter binding are summarized in Table 2.2.

<table>
<thead>
<tr>
<th></th>
<th>wt AR (kcal/mol)</th>
<th>T877A (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta G_{\text{total}}$</td>
<td>$-16.9 \pm 0.1$</td>
<td>$-16.4 \pm 0.1$</td>
</tr>
<tr>
<td>$\Delta G_c$</td>
<td>$-2.3 \pm 0.1$</td>
<td>$-2.5 \pm 0.2$</td>
</tr>
<tr>
<td>SD of fit$^b$</td>
<td>0.043</td>
<td>0.060</td>
</tr>
</tbody>
</table>

$^a$Values calculated using the relationship $\Delta G_i = -RT \ln(k_i)$. Errors represent one SD reported from the program Scientist.

$^b$SD of global fit in units of fractional saturation.

Discussion

Hydrodynamic and thermodynamic properties of human androgen receptor

Here we have examined the physical and functional attributes of human AR using a rigorous and quantitative approach and under a standardized set of conditions. In doing so, we hope to gain insight into the quantitative mechanisms by which steroid receptors generate receptor-specific gene regulation. We first discovered that both wt AR and T877A exist only as monomers, showing no evidence of dimerization up to micromolar concentrations. We note that this result is at odds with the traditional model of receptor function$^3$, and with early semi-quantitative

34
studies indicating that AR formed disulfide-linked dimers\textsuperscript{51}. We still detect a small amount (5-10\% of an irreversible aggregate that corresponds to AR trimer. Notably, we also observe this species using a non-equilibrium assay (native gel electrophoresis) similar to one first used to report the existence of disulfide-linked dimers\textsuperscript{51}. Indeed, we find that if DTT levels are decreased to less than 1 mM, the percentage of this aggregate increases significantly, indicating that it is disulfide-linked (data not shown). Unfortunately, decreased DTT levels also increase AR polydispersity, making it intractable for rigorous analysis and interpretation. (Freeze-thawing of purified AR generated a similar result, presumably by decreasing DTT activity within the buffer). Taken together, we therefore postulate that the irreversible AR aggregate observed here is functionally irrelevant, and that the monomer should be considered as the active species.

Although the finding that AR does not readily dimerize contradicts previous interpretations, it is fully consistent with crystallographic work demonstrating that the isolated hormone-binding domain – the domain thought to mediate dimerization for other receptors – crystallizes as a monomer\textsuperscript{64}. Our findings are also consistent with genome-wide studies indicating that AR, much more so than other receptors, preferentially binds to and transcriptionally activates at half-sites rather than palindromes\textsuperscript{72}. Thus under the conditions examined here, AR shows no evidence of reversible self-association. However, processes such as DNA binding, presence of accessory proteins, or molecular crowding events could result in self-association via thermodynamic coupling. With regard to the strong AR cooperativity observed here, a number of biochemical studies have indicated that the receptor is capable of cooperative interactions\textsuperscript{55,73,74}, thus confirming the results seen here. However, differences in experimental approach, receptor preparation and data analysis make it difficult to directly compare those finding to the values reported here. Finally, as seen in Table 2.2, wt AR and T877A share essentially identical energetic parameters, suggesting that interaction energetics
do not play a role in explaining the unique functional properties of the mutant\textsuperscript{64}. However, as evidenced by our 100 mM NaCl sedimentation equilibrium studies, T877A is more stable than wt. This could suggest that another mechanism by which T877A contributes to prostate cancer is via enhanced stability, thus leading to inappropriate persistence of function. This possibility is currently being investigated.

**Energetic comparison to other receptors**

As outlined in the introduction, we are attempting to systematically dissect the energetics of steroid receptor-promoter interactions under a rigorously controlled set of conditions. For receptors such as ER-\(\alpha\) and the two PR isoforms, this corresponds to the traditional understanding that monomers dimerize in the absence of DNA (\(k_{\text{dim}}\)) and assemble as pre-formed dimers capable of engaging in inter-site cooperativity (\(k_c\)). These values are plotted in Figure 2.7. Also shown are the cooperativity terms resolved for wt AR, T877A and GR. However, since these latter receptors do not show any evidence for dimerization, we can only plot lower limits as estimated from our analytical ultracentrifugation studies. For example, noting that our studies should be able to easily detect 10\% dimer, and the highest wt AR concentration we examined was 1.4 µM, this places a lower limit for dimerization (as emphasized by the downward arrow) at 14 µM. A similar approach was used for T877A and GR. Note that all parameters are plotted in log units on the y-axis.

Despite being a homologous family of transcription factors, the results in Figure 2.7 make clear that steroid receptor dimerization and cooperative energetics vary enormously. In particular, dimerization energetics (\(k_{\text{dim}}\)) are generally inversely proportional to cooperative energetics (\(k_c\)). Thus, ER-\(\alpha\) displays a sub-nanomolar dimerization affinity but generates statistically non-existent cooperativity. At the opposite extreme are the AR proteins and GR,
which have dimerization limits that are four to five orders of magnitude weaker than ER-α, and cooperativity values two orders of magnitude greater. The second observation is that the inverse relationship between dimerization and cooperativity trends along evolutionary lines. Thus PR-B, wt AR, T887A and GR – all closely related, subgroup 3C receptors – partition their dimerization and cooperative energetics similarly. In sharp contrast, ER-α, the more distantly related subgroup 3A receptor, maintains a distinct distribution. (It is also clear that PR-A, a subgroup 3C receptor, is not fully consistent with this argument since it exhibits weak cooperativity. We note that PR-A is still capable of generating 1000-fold cooperative stabilization on a different promoter architecture. We find similar promoter-specific differences in cooperativity for other subgroup 3C receptors via manipulation of HRE phasing and periodicity. Moreover, the extent of cooperativity now varies between these receptors, suggesting a critical role for cooperativity in generating receptor-specific transcriptional responses. By contrast, ER-α cooperativity is not detectable on seven promoter architectures we have tested to date.

The consequences of such large differences in energetics can be better appreciated in Figure 2.8. Shown are the probabilities of observing the fully ligated HRE2 promoter (i.e. the presumed transcriptionally active microstate) as a function of total receptor concentration under our standardized conditions. Overlaid are the probabilities of observing receptor dimer, calculated from either an experimentally determined dimerization constant (ER-α, PR-A and PR-B) or an estimated lower limit (wt AR, T877A and GR). All probabilities were calculated using the energetics shown in Figure 2.7 and equations (2.9) and (2.10). It is immediately obvious that for the subgroup 3A receptor, ER-α, a significant dimer population exists over the entire concentration range in which receptor binds to the promoter. (Interestingly, the strong dimerization energetics of ER-α predict an abundance of functionally active full-sites rather than
half-sites throughout the genome, which indeed is the case\cite{72}.) By contrast, for all subgroup 3C receptors, little dimer exists from low to full DNA saturation. Although these results do not allow us to address whether DNA binding occurs via a pre-formed dimer or successive monomer assembly reaction, they nonetheless predict a deficiency of subgroup 3C dimers under \textit{in vitro} DNA binding conditions and raise questions about the relevance (or existence) of such dimers in cellular function.

\textbf{Figure 2.7.} Microstate energetics of steroid receptor dimer assembly at the HRE$_2$ promoter. Circles represent receptor dimerization affinity ($k_{\text{dim}}$) and triangles represent inter-site cooperativity ($k_c$). Since dimerization was not observed for wt AR, T877A and GR, downward arrows have been added to indicate that plotted values represent lower limits. Error bars represent 67% confidence intervals.
Figure 2.8. Probabilities of receptor dimer formation and fully-ligated HRE\textsubscript{2} promoter microspecies under standard state conditions. Probabilities determined from experimentally measured energetics or lower limit estimates. Blue lines represent the probabilities of receptor dimers and red lines represent the probabilities of the fully-ligated HRE\textsubscript{2} promoter. Because dimers have not been observed for wt AR, T877A and GR, blue lines for these proteins represent limits.

*Has evolution tinkered with the functional energetics of steroid receptors?*

Our results demonstrate that steroid receptor family members differentially partition their microscopic interaction energetics, and do so by multiple orders of magnitude. Furthermore, the manner in which energetics are partitioned mirrors the evolutionary divergence of the receptor family. Noting that all the receptors descend from an ER-like ancestral protein, it may be the case that the 3C subgroup receptors (AR, GR, MR and PR) lost the ability to strongly dimerize but gained cooperative binding. This may seem unlikely since sequence-specific transcription factors
typically maintain strong dimerization as a tool for increasing DNA binding affinity and specificity. However, a redistribution of interaction energetics may serve a more important role by generating receptor-specific promoter occupancy. For example, simulations\textsuperscript{32} indicate that if a loss in receptor dimerization is compensated by inter-site cooperativity, it is not only possible to maintain promoter occupancy, but also preferential promoter occupancy – that is, occupancy of a particular receptor even in the presence of competitive binding by other receptors with similar DNA binding affinities. Second, experimental studies demonstrate that it is receptor monomers rather than (the previously assumed) dimers that assemble at clusters of half-sites, and that monomer binding is highly cooperative\textsuperscript{75}. Moreover, multiple types of cooperativity exist – between pairs of monomers, monomers and dimers, and the pairs of dimers seen here. Taken together, these results suggest that receptor assembly-state, promoter binding stoichiometry, and the number and layout of promoter binding sites are critical factors for specifying receptor occupancy at a promoter.

In summary, our findings suggest that nature has selected for receptor-specific differences in energetics as a basis for generating receptor-specific transcriptional activation. Specifically, the combination of different interaction energetics and different promoter architectures may serve as a combinatorial framework for generating receptor-specific promoter occupancy, and thus receptor-specific transcriptional activity. This is an appealing hypothesis for explaining how receptors such as AR, GR, MR and PR can bind identical response elements yet regulate different subsets of genes\textsuperscript{77–79}. Moreover, it is consistent with the recent finding that despite the apparent complexity of gene regulation, DNA binding energetics are the dominant factor in controlling transcriptional activity at different response elements\textsuperscript{62}. Finally, such a hypothesis implies that steroid receptors are energetically poised to carry out function, and predicts a multidimensional “binding affinity landscape” for each receptor\textsuperscript{80}. 
Materials and Methods

Expression and purification of full-length human AR proteins

An expression plasmid encoding the wild-type version of full-length human AR (residues 1-919) was donated by Dr. Elizabeth Wilson (University of North Carolina, Chapel Hill). The AR-encoding region was PCR amplified and cloned into a pBAC baculovirus vector (EMD, formerly Novagen) encoding an N-terminal hexahistidine tag. Site-directed mutagenesis was then used to create the T877A mutation. Both receptors were expressed in baculovirus-infected Sf9 insect cells using a multiplicity of infection of 1. 24 hour post-infection, cells were treated with 1 µM of the synthetic androgen R1881. Cells were harvested 24 h later.

All purification steps were carried out at 4 °C and in the presence of 10 µM R1881. Cells were Dounce-homogenized in a buffer containing 20 mM Tris (pH 8.0 at 4 °C), 10 % glycerol, 500 mM NaCl, 10 mM β-ME, 10 µM R1881, 25 mM imidazole and protease inhibitors (Complete, EDTA-free, Roche). Following centrifugation, the AR-containing supernatant (either wild-type or T877A) was incubated for 1 h with Ni-NTA agarose (Qiagen). The AR-bound resin was washed extensively with homogenation buffer and receptor was eluted using the same buffer now containing 250 mM imidazole. Eluted receptor was then chromatographed on a Sephacryl S-400 size exclusion column (GE Healthcare) equilibrated in 20 mM Tris (pH 8.0 at 4 °C), 500 mM or 100 mM NaCl, 1 mM CaCl$_2$, 2.5 mM MgCl$_2$, 1 mM DTT and 10 µM R1881. The identity of wt AR and T877A were confirmed by MALDI-TOF mass spectrometry, and concentrations were determined using a calculated extinction coefficient of 93,170 M$^{-1}$ cm$^{-1}$. Finally, because we found that even a single freeze-thaw cycle generated AR aggregation and polydispersity, only freshly purified protein was used in the experiments shown here.
**Sedimentation velocity**

Sedimentation velocity experiments were carried out using a Beckman XL-A analytical ultracentrifuge equipped with absorbance optics, using two-sector epon centerpieces and an An-50 Ti rotor. Each AR protein was sedimented in a buffer containing 20 mM Tris (pH 8.0 at 4 °C), 1 mM CaCl$_2$, 2.5 mM MgCl$_2$, 1 mM DTT, 10 µM R1881 and NaCl concentrations of either 100 or 500 mM. Other than steroid type, these conditions were identical to our previous work on GR, the two PR isoforms and ER-α$^{31,32,34,35}$. Receptors were sedimented at concentrations ranging from 0.11 to 1.5 µM at 50,000 rpm and 4 °C. Samples were monitored at 230 nm and scans were collected as quickly as the instrument would allow. Sedimentation coefficient ($s$) distributions were calculated using the program Sedfit$^6$ and corrected to standard conditions (water at 20 °C)$^9$. The frictional coefficient ($f$) was calculated using the Svedberg equation:

$$f = \frac{M(1 - \bar{v}\rho)}{N_A s}$$

(2.1)

Where $M$ is the calculated AR monomer molecular mass, $\bar{v}$ is the AR partial specific volume$^8$, $\rho$ is the solvent density$^9$ and $N_A$ is Avogadro’s number. The density was calculated from the buffer composition and temperature$^{40}$, and the partial specific volume was calculated by summing the partial specific volumes of the individual amino acids (0.7171 mL/g for both receptors)$^8$.

**Sedimentation equilibrium**

For T877A, sedimentation equilibrium studies were carried out under the identical 100 and 500 mM NaCl buffer conditions as used in the velocity studies. For wt AR, studies were carried out only at 500 mM NaCl since the receptor slowly and irreversibly aggregates after ~24 hours under low salt conditions. All samples were allowed to reach sedimentation equilibrium using six-channel Epon centerpieces. Wild-type AR was loaded at three concentrations, 0.3, 0.4 and 0.6
µM. T877A was loaded at 0.3, 0.5 and 1.3 µM. Samples were equilibrated at 14,000, 17,000 and 21,000 rpm and judged to be at equilibrium by successive subtraction of scans. Data sets were analyzed individually and globally using nonlinear least squares analysis as implemented in the program Scientist 3.0 (Micromath). Models were constructed using the following equation to resolve the reduced molecular mass ($\sigma_i$):

$$Y_r = \delta + \sum_{i=1}^{n} \alpha_i \exp \left( \sigma_i \frac{r^2 - r_0^2}{2} \right)$$  \hspace{1cm} (2.2)

Where $Y_r$ is the absorbance at radius $r$, $\delta$ is the baseline offset, $n$ is the number of non-interacting species and $\alpha_i$ is the absorbance of the $i$th species at the reference radius, $r_0$. The reduced molecular mass of the $i$th species is defined as:

$$\sigma_i = \frac{M(1 - \bar{v} \rho) \omega^2}{RT}$$  \hspace{1cm} (2.3)

Where $M$ is the weight-average molecular mass, $\bar{v}$ is the partial specific volume of AR, $\rho$ is the solvent density, $\omega$ is the angular velocity, $R$ is the gas constant and $T$ is the absolute temperature.$^{40}$

**DNA preparation for DNase I footprint titrations**

A promoter vector encoding two tandemly linked hormone response elements (HRE$_2$; see Figure 2.1.b) was donated by Dr Kathryn Horwitz (University of Colorado Anschutz Medical Campus). Each HRE corresponds to an imperfect palindrome derived from the tyrosine aminotransferase promoter, AGAACAaggTGTACA$^{84}$, spaced 25 bp apart. A reduced valency template (HRE$_1$) containing a C-to-A point mutation in each half site of the distal HRE (designated as site 1) was created in-house. Each template was excised from its respective vector using HindIII and AatII to generate a 1 kb promoter fragment and end-labeled with $^{32}$P using a Klenow fill-in reaction. The
proximal HRE of each fragment (site 2) was positioned 100 bp from the 3’ end of the labeled strand. The HRE₂ promoter is identical to that of the previously named GRE₂ and PRE₂ promoters used in our earlier work on GR and the two PR isoforms²⁹,³⁰,³². The promoter differs only in response element sequence when compared to the ERE₂ template used in our work on ER-α³¹.

**Individual-site binding experiments**

DNA-binding experiments were conducted using quantitative DNase I footprint titrations as described by Ackers and co-workers with minor modifications⁴¹–⁴³. All reactions were carried out in a buffer containing 20 mM Tris (pH 8.0 at 4 °C), 100 mM NaCl, 1 mM CaCl₂, 2.5 mM MgCl₂, 1 mM DTT, 10 µM R1881, 100 µg/mL BSA and 2 µg/mL salmon sperm DNA. Other than BSA and salmon sperm DNA, this buffer is identical to that used in our 100 mM NaCl sedimentation studies. Each reaction mixture contained 15,000 cpm of freshly labeled HRE₁ or HRE₂ promoter. Freshly prepared AR (either wild-type or T877A) was added to each reaction sample, covering a concentration range from nanomolar to micromolar. Samples were allowed to equilibrate at 4 °C for at least 1 h. Promoter DNA concentrations (~10 pM) were estimated to be well below AR DNA binding affinity, justifying the assumption that AR_free ≈ AR_total. Denatured DNA fragments were electrophoresed on 6% acrylamide-urea gels and visualized using phosphorimaging. Band densities were determined using the program ImageQuant TL 7.0 (GE Healthcare). Individual-site binding isotherms were calculated as described previously⁴².

**Resolution of microscopic interaction free energies**

Footprint titrations resolve the fractional occupancy of binding \( \bar{Y} \) at each HRE. The statistical thermodynamic expressions that describe the resultant individual-site binding isotherms are constructed by summing the probabilities of each microstate that contributes to binding at each
site. A detailed description for constructing these expressions has been presented elsewhere. In short, the probability \( f_s \) of each microstate is defined as:

\[
f_s = \frac{e^{(-\Delta G_s/RT)} x^j}{\sum_{s=1}^j e^{(-\Delta G_s/RT)} x^j}
\]  

(2.4)

where \( \Delta G_s \) is the free energy of configuration of energy state \( s \) relative to that of the unliganded reference state, \( x \) is the free AR monomer concentration, \( j \) is the stoichiometry of AR monomers bound to a response element, \( R \) is the gas constant and \( T \) is the absolute temperature. The relationship between each free energy change and its association constant is described by the equation \( \Delta G_i = -RT \ln k_i \).

Isotherms were analyzed to resolve the total binding affinity \( (K_{tot}) \) for assembling two AR monomers at a palindromic HRE, and the inter-site cooperativity \( (k_c) \) between AR dimers bound at separate HREs (see Figures 2.1.b and 2.1.c). Equation (2.5) describes AR binding by such a model to an HRE\(_1\) isotherm:

\[
\bar{Y}_{HRE_1} = \frac{K_{tot} x^2}{1 + K_{tot} x^2}
\]

(2.5)

where \( x \) is the free AR monomer concentration. Using a similar approach, equation (2.6) describes AR binding to each isotherm of the HRE\(_2\) promoter:

\[
\bar{Y}_{HRE_2} = \frac{K_{tot} x^2 + K_c^2 k_c x^4}{1 + 2K_{tot} x^2 + K_c^2 k_c x^4}
\]

(2.6)

To resolve \( K_{tot} \) and \( k_c \), the isotherms from all footprint titrations were analyzed globally using Scientist 3.0 (Micromath). The apparent binding affinity \( (K_{app}) \) was determined by fitting the HRE\(_1\) isotherms to the Langmuir binding equation:

\[
\bar{Y} = \frac{K_{app} x}{1 + K_{app} x}
\]

(2.7)
Finally, since receptor binding at specific DNA sites does not result in complete protection from DNase, all data were treated as transition curves \( \overline{Y}_{\text{app}} \) fit to upper \((m)\) and lower \((b)\) end points:

\[
\overline{Y}_{\text{app}} = (m - b) \overline{Y} + b
\]

\[ \text{(2.8)} \]

**Probabilities of HRE\(_{2}\) occupancy and dimer formation for the steroid receptors.**

Figure 2.8 presents the probability of HRE\(_{2}\) promoter occupancy and dimer formation for steroid receptor family members. The probability of observing a fully-ligated HRE\(_{2}\) promoter for each steroid receptor \(P_{\text{HRE2-SR}}\) was calculated as follows:

\[
P_{\text{HRE2-SR}} = \frac{K_{\text{tot-SR}}^2 k_{c-SR} x_{SR}^4}{1 + 2 K_{\text{tot-SR}} x_{SR}^2 + K_{\text{tot-SR}}^2 k_{c-SR} x_{SR}^4}
\]

\[ \text{(2.9)} \]

Where \(K_{\text{tot-SR}}\) is the total affinity of two monomers at a HRE, \(k_{c-SR}\) is the intersite cooperativity and \(x_{SR}\) is the total receptor concentration (in monomer units). The probability of observing a steroid receptor dimer \(P_{\text{dimer-SR}}\) was calculated for each receptor using the following expression:

\[
P_{\text{dimer-SR}} = \frac{x_{d-SR}}{x_{SR}} = \frac{x_{SR} - x_{m-SR}}{x_{SR}}
\]

\[ \text{(2.10)} \]

Where \(x_{d-SR}\) is the concentration of a steroid receptor dimer (in monomer units), \(x_{SR}\) is the total steroid receptor concentration and \(x_{m-SR}\) is the monomer concentration of a steroid receptor:

\[
x_{m-SR} = \frac{-1 + \sqrt{1 + 8 k_{\text{dim-SR}} x_{SR}}}{4 k_{\text{dim-SR}}}
\]

\[ \text{(2.11)} \]

Where \(k_{\text{dim-SR}}\) is the dimerization equilibrium constant. Since dimers were not observed for GR\(^9\), wild type AR and T877A, the lower limit values based on sedimentation velocity and equilibrium analyses were used (see text for further details).
CHAPTER III

GLUCOCORTICOID RECEPTOR-DNA DISSOCIATION KINETICS MEASURED IN VITRO REVEAL EXCHANGE ON THE SECOND TIMESCALE

Summary

The glucocorticoid receptor (GR) is a member of the steroid receptor family of ligand-activated transcription factors. Recent live cell imaging studies have revealed that GR interactions with chromatin are highly dynamic, with average receptor residence times of only seconds. These findings were surprising since early kinetic studies found that GR-DNA interactions in vitro were much slower, having calculated residence times of minutes to hours. However, these latter analyses were carried out at a time when it was only possible to work with either partially purified holoreceptor or its purified but isolated DNA binding domain. Noting these limitations, we reexamined GR-DNA dissociation kinetics using highly purified holoreceptor shown amenable to rigorous study. We first observe that GR-DNA interactions in vitro are not slow as previously thought, but converge with in vivo behavior, having residence times of only seconds to tens of seconds. This rapid exchange is seen at six individual response elements and the multisite MMTV promoter used in live cell imaging. Second, GR dissociation rates are identical for all response elements. Thus previously observed differences in receptor affinity toward these sequences are not due to differences in off-rate but in on-rate. Finally, dissociation kinetics are biphasic in character. A minimal kinetic model consistent with the data is that DNA-bound GR interconverts between states on a second timescale, with dissociation occurring via a multistep process. We speculate that receptor interconversion in this timeframe is recognizable by the coregulatory proteins that interact with GR, leading to unique transcriptional responses.

2 At the time of submission of this thesis, Chapter III was accepted by the journal Biochemistry (DOI: 10.1021/acs.biochem.5b00693).
Introduction

The glucocorticoid receptor (GR) is a member of the steroid receptor family of ligand-activated transcription factors\(^3\). Upon binding ligand, the receptor activates gene expression by first assembling at hormone response elements (HREs), typically as a dimer. Receptor-DNA binding is coupled to recruitment of coregulatory proteins, chromatin remodeling, and transcriptional activation. Advances in live cell imaging have allowed visualization of a subset of these events in vivo, with the finding that GR interactions with chromatin are highly dynamic. For example, initial studies using an engineered array of the mouse mammary tumor virus (MMTV) promoter revealed an average GR residence time of only 10 sec\(^57\). More recent studies of receptor interactions at single-copy, endogenous promoters have also reported fast exchange\(^58\).

Rapid exchange in vivo was surprising since early in vitro work had demonstrated that GR-DNA interactions were much slower. For instance, kinetic dissociation studies using full-length GR and the MMTV promoter resolved an off-rate of \(1.1 \times 10^{-4} \text{ s}^{-1}\), corresponding to an average residence time of 2.5 hr\(^60\). Similar results were seen using only the GR DNA-binding domain (DBD) and a single HRE, resolving an off-rate of \(2.2 \times 10^{-4} \text{ s}^{-1}\) or average residence time of 1.3 hr\(^61\). This discrepancy between in vitro and in vivo GR binding dynamics led to the proposal that in live cells, the receptor is actively displaced from DNA as a part of the chromatin remodeling process\(^86\).

Although the above in vitro studies were once state of the art, they were also carried out at a time when it was only possible to study either partially purified holoreceptor, or the purified but isolated DBD. The advent of baculovirus-insect cell expression systems has now allowed high-yield expression and purification of intact GR, which in turn has facilitated more detailed biophysical investigations. These studies have revealed that the highly purified holoreceptor displays a number of attributes not seen with unpurified GR or its isolated
domains. These include differences in receptor assembly-state, DNA binding energetics, cooperative assembly at complex promoters, and the role of such interactions in regulating transcription in vivo\(^{32,62}\). In light of these developments, and noting that live cell imaging studies are not yet positioned to gain physical mechanisms of receptor-DNA interactions, we decided to reexamine in vitro GR dissociation kinetics. Our goal was to shed additional light on the kinetic mechanisms by which the GR holoprotein interacts with its response elements and multisite promoters. As shown in Figure 3.1a, we analyzed a series of six well-characterized response elements\(^{24,62}\), and the MMTV promoter used in live cell imaging (Figure 3.1b)\(^{57}\).

![Figure 3.1](image.png)

**Figure 3.1.** Schematics for six response elements and the mouse mammary tumor virus (MMTV) promoter. (a) Sequences for the six palindromic-like HREs. Half-sites are represented by black rectangles; arrows indicate relative orientation. (b) Response elements located in the MMTV promoter. Arrows below schematic represent relative orientation. Numbers represent the upstream distance from transcriptional start site, noted as arrow above schematic. Site 1 represents sequence GTTACAaaaTGTCT. Sites 2 through 4 represent half-sites with sequence TGTCT.
Our studies reveal that GR-DNA interactions in vitro are not slow as previously thought, but converge with in vivo behavior, with GR residence times of only seconds to tens of seconds. Rapid exchange is observed at all response elements and the MMTV promoter. We also find that GR dissociation rates are identical for all response elements tested, indicating that previously observed differences in receptor binding affinity toward these sequences are not due to differences in off-rate but in on-rate. Finally, we observe that GR-DNA dissociation kinetics are biphasic in character. A minimal kinetic model consistent with all data is that the DNA-bound receptor interconverts between states on a second timescale, resulting in a multistep dissociation process. We speculate that interconversion within this timeframe is recognizable by the coregulatory proteins that interact with GR, leading to unique transcriptional responses.

Results

Full-length GR is amenable to kinetic analysis

We recently developed protocols for high-yield expression and purification of full-length, human GR\textsuperscript{32,62}. We further showed that the purified receptor is structurally and functionally homogeneous, and therefore amenable to thermodynamic studies. Here we reproduce a portion of these findings to lend credence to our kinetic investigations. As shown in Figure 3.2a, the full-length receptor can be purified to greater than 90% as judged by densitometry. As shown in Figure 3.2b, sedimentation velocity studies indicate that GR sediments primarily as a 4.2 S species with a molecular mass of 90 kDa. This agrees with the calculated mass of the GR
monomer (90,925 Da), and is consistent with our more comprehensive sedimentation studies demonstrating that GR shows no evidence of self-association up to and above micromolar concentrations. (The remaining species from 5 to 10 S reflect small amounts of irreversible GR aggregates we believe to be functionally inactive)\textsuperscript{32}. Shown in Figure 3.2c is an equilibrium-binding isotherm for GR assembly at the TAT\textsubscript{3} response element. The total binding affinity was determined to be $1.1 \pm 0.2 \times 10^{13} \text{M}^{-2}$ or $-16.5 \text{kcal/mol}$, indicating that GR binds with strong affinity. Separate fitting to a Hill equation resolved a Hill coefficient of $1.8 \pm 0.2$, indicating substantial cooperativity between the two bound monomers. All of these findings are statistically identical to our previous studies using both His-tagged and FLAG-tagged GR\textsuperscript{32,62}, and suggest that the receptor purified here is amenable to detailed kinetic analysis.

\textbf{Figure 3.2.} Purification and biophysical characterization of full-length human GR. (a) Coomassie-stained SDS-PAGE image of 5 µg purified GR. (b) Sedimentation coefficient distribution of 0.5 µM GR determined by Sedfit analysis.\textsuperscript{67} (c) Fractional saturation ($\bar{Y}$) of the TAT\textsubscript{3} response element by GR determined by quantitative equilibrium footprinting. Solid line represents best fit to a contracted Adair equation\textsuperscript{32}.

\textit{GR-response element dissociation kinetics are biphasic and occur on the second timescale}

We first studied GR dissociation kinetics at the TAT\textsubscript{3} response element. Shown in Figure 3.3a is a kinetic footprint showing GR dissociation from this sequence over 600 s. The resultant decay
curve is shown in Figure 3.3b, with the first 50 s shown in the inset. Visual inspection indicates that the half-life for the GR-DNA complex is between 10 to 20 s, comparable to the short residence time seen in live cells. Fitting to a single exponential decay model (dotted line) resolved an observed rate constant, $k_{obs}$, of 0.025 s$^{-1}$. However, the single exponential model poorly describes the data, particularly in the first 50 sec of dissociation (Figure 3.3b inset). We therefore fit the data to a double-exponential or biphasic decay model, which resulted in a visually improved fit over the entire time course (solid line). An F-test, which accounts for additional fitting parameters in the biphasic model, confirmed that the fit is statistically improved over the single exponential fit. The biphasic fit resolved a fast phase rate constant ($k_{obs,1}$) equal to 0.3 ± 0.1 s$^{-1}$ and a slow phase rate constant ($k_{obs,2}$) equal to 0.010 ± 0.002 s$^{-1}$. These values translate to average GR-DNA residence times of 4 and 100 sec, respectively. Finally, the amplitude of each phase was comparable, with a 40 % contribution from the fast phase and a 60 % contribution from the slow phase.

To further probe the basis of GR-DNA dissociation, we repeated the above experiment using GR concentrations ranging from 0.14 to 1.4 µM. These concentrations correspond to equilibrium GR-DNA occupancies ranging from 0.2 to over 0.9 fractional saturation units (see Figure 3.2c). All time courses were visually similar to that in Figure 3.3b (not shown), and fitting of the biphasic model again resulted in statistically improved fits. As shown in Figure 3.4a and Table 3.1, the resultant amplitude terms (plotted as a ratio of A1/(A1 + A2)), show no evidence of GR concentration dependence. A similar result is seen for the two observable rate constants (Figure 3.4b and Table 3.1). This lack of concentration dependence indicates that only a single ligation species is present on the DNA regardless of GR concentration. This is the predicted result based on the strong cooperativity seen for GR-DNA binding in the equilibrium studies, and suggests that only a GR dimer exists on the DNA in the kinetic studies.
Figure 3.3. Kinetic footprint and decay curve for GR dissociation from the TAT$_3$ response element. (a) Representative autoradiogram of GR dissociation from TAT$_3$ sequence. GR concentration was 1.4 μM. Position of binding site is indicated by schematic at left. Exposure time to unlabeled competitor DNA increases from left to right as indicated by triangle. (b) Decay curve determined from analysis of footprint image in (a). Open circles represent time-dependent GR occupancy at TAT$_3$ response element. Dotted line represents best fit to a single-phase exponential decay model. Solid line represents best fit to a biphasic decay model. Shown in inset is fit covering the first 50 s of GR dissociation.
Figure 3.4. Relative amplitudes and observed rate constants for concentration-dependent GR dissociation from the TAT$_3$ response element. (a) Open circles represent amplitude of the fast phase, $A_1$, expressed as a ratio of $A_1/(A_1 + A_2)$. Error bars represent standard deviation propagated from error in both amplitude terms. Dashed line represents weighted-averaged amplitude across all GR concentrations ($0.5 \pm 0.1$). (b) Observed rate constants and standard deviations for $k_{\text{obs},1}$ and $k_{\text{obs},2}$ as determined from biphasic decay model. Dashed lines represent the weighted averages across all GR concentrations, calculated to be $0.3 \pm 0.2$ s$^{-1}$ and $0.014 \pm 0.006$ s$^{-1}$, respectively.
Table 3.1. Observed rate constants and relative amplitudes for biphasic dissociation of GR from the TAT₃ response element*

<table>
<thead>
<tr>
<th>[GR] (µM)</th>
<th>kₜₗ,₁ (s⁻¹)</th>
<th>kₜₗ,₂ (s⁻¹)</th>
<th>A₁ / (A₁+A₂)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.14</td>
<td>1.2 (-2., +2)</td>
<td>0.029 ± 0.008</td>
<td>0.3 ± 0.2</td>
</tr>
<tr>
<td>0.30</td>
<td>1.7 (-1.7, +2.5)</td>
<td>0.047 ± 0.01</td>
<td>0.5 ± 0.2</td>
</tr>
<tr>
<td>0.47</td>
<td>0.23 ± 0.08</td>
<td>0.007 ± 0.006</td>
<td>0.6 ± 0.2</td>
</tr>
<tr>
<td>0.50</td>
<td>0.9 ± 0.5</td>
<td>0.02 ± 0.01</td>
<td>0.6 ± 0.3</td>
</tr>
<tr>
<td>0.50</td>
<td>1.0 ± 0.4</td>
<td>0.018 ± 0.007</td>
<td>0.7 ± 0.2</td>
</tr>
<tr>
<td>0.60</td>
<td>0.2 ± 0.1</td>
<td>0.02 ± 0.01</td>
<td>0.5 ± 0.4</td>
</tr>
<tr>
<td>0.60</td>
<td>0.5 ± 0.5</td>
<td>0.017 ± 0.01</td>
<td>0.6 ± 0.4</td>
</tr>
<tr>
<td>0.73</td>
<td>0.7 ± 0.2</td>
<td>0.014 ± 0.003</td>
<td>0.5 ± 0.1</td>
</tr>
<tr>
<td>1.20</td>
<td>0.9 ± 0.2</td>
<td>0.029 ± 0.007</td>
<td>0.70 ± 0.09</td>
</tr>
<tr>
<td>1.40</td>
<td>0.28 ± 0.07</td>
<td>0.010 ± 0.002</td>
<td>0.4 ± 0.1</td>
</tr>
<tr>
<td>1.40</td>
<td>0.52 ± 0.1</td>
<td>0.014 ± 0.003</td>
<td>0.5 ± 0.1</td>
</tr>
</tbody>
</table>

*Parameters were resolved by individually fitting each progress curve (n=1) to a two-phase exponential decay model using the program, Scientist. Errors represented 68% confidence limits determined via fitting.

GR-DNA dissociation kinetics are independent of HRE sequence

GR binds to a variety of HREs and with a wide range of affinities²⁶,⁶². To determine if the results seen for TAT₃ were recapitulated on other HREs, and to address the kinetic basis for differences in GR binding affinity at these sequences, we analyzed receptor dissociation kinetics at the five remaining HREs shown in Figure 3.1a. The resultant decay curves are shown in Figure 3.5.

Following our approach for analyzing the TAT₃ sequence, we again fit each dissociation curve to either a single or double exponential decay model. We found that for all HREs, the biphasic decay model resulted in a statistically improved fit. The resultant fast and slow phase rate constants, kₜₗ,₁ and kₜₗ,₂, for all sequences examined are shown in Figure 3.6 and Table 3.2.

Surprisingly, we see that neither rate constant changes as function of GR binding affinity, even though affinity at these sequences covers an 80-fold range⁶². Thus GR-DNA-binding affinity at these sequences is not controlled by differences in off-rate, but instead by on-rate.
Figure 3.5. Decay curves for GR dissociation from five response elements. Open circles represent time-dependent GR occupancy at each response element. Dotted line represents best fit to single-exponential model; solid line represents fit to biphasic decay model. GR concentrations were 0.5 μM for Pal, 0.7 μM for CGT, 0.5 μM for SGK, 0.6 μM for Cons, and 1.1 μM for FKBP5.
Figure 3.6. Observed rate constants for GR dissociation from six response elements. Open circles represent either $k_{\text{obs},1}$ or $k_{\text{obs},2}$ determined for each sequence by global analysis of at least two decay curves. Error bars represent standard deviation. GR binding affinity for the six sequences decreases 80-fold from left to right. Dashed lines represent the weighted average for each rate constant across all response elements. Averages were calculated to be 0.5 ± 0.2 s$^{-1}$ for $k_{\text{obs},1}$ and 0.020 ± 0.006 s$^{-1}$ for $k_{\text{obs},2}$. 
Table 3.2. Observed rate constants for biphasic dissociation of GR from six individual response elements*

<table>
<thead>
<tr>
<th>HRE</th>
<th>$k_{obs,1}$ (s$^{-1}$)</th>
<th>$k_{obs,2}$ (s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pal</td>
<td>2.1 (-2.1, +4.1)</td>
<td>0.04 ± 0.01</td>
</tr>
<tr>
<td>CGT</td>
<td>1.0 ± 0.3</td>
<td>0.019 ± 0.003</td>
</tr>
<tr>
<td>TAT$_3$</td>
<td>0.4 ± 0.1</td>
<td>0.019 ± 0.003</td>
</tr>
<tr>
<td>SGK</td>
<td>0.9 ± 0.3</td>
<td>0.018 ± 0.006</td>
</tr>
<tr>
<td>Cons</td>
<td>0.8 ± 0.4</td>
<td>0.05 ± 0.03</td>
</tr>
<tr>
<td>FKBP5</td>
<td>1.6 ± 0.9</td>
<td>0.036 ± 0.007</td>
</tr>
</tbody>
</table>

*Parameters for each sequence were resolved by globally fitting two to eleven decay curves, covering a range of GR concentrations, to a two-phase exponential decay model using the program, Scientist. GR concentrations are as noted in Fig. 3.5. Errors represented 68% confidence limits determined via fitting.

GR exchange with the MMTV promoter also occurs on the second timescale

We next examined GR dissociation kinetics from the MMTV promoter used in live cell imaging studies. As shown in Figure 3.1b, this promoter contains one imperfect palindrome (HRE 1) and three recognizable half-sites (HREs 2-4). The GR binding stoichiometry at each site has not been definitively established, although early studies suggested that dimers bound at all sites$^{60}$. Shown in Figure 3.7a is the decay curve for GR dissociation from the imperfect palindrome (HRE 1), and shown in Figure 3.7b is the curve for its nearest half-site (HRE 2). Both decay curves were analyzed using single- and double-exponential decay models. On the basis of visual inspection and an F-test, we again find that GR dissociation at each site is statistically best described using a biphasic decay model (solid lines in Figure 3.7a, b). For site 1, we resolved a fast phase rate constant of 0.9 ± 0.2 s$^{-1}$ and a slow phase rate constant of 0.040 ± 0.009 s$^{-1}$. Comparable values were seen for site 2, with fast and slow rate constants of 1.0 ± 0.3 s$^{-1}$ and 0.025 ± 0.008 s$^{-1}$, respectively. Although quantification was not reliable for sites 3 and 4 due to background noise, visual inspection indicated that GR dissociation rates were similar to those observed for sites 1
and 2. In sum, GR dissociation from the MMTV promoter follows similar kinetics to those seen for individual HREs.

Figure 3.7. GR dissociation from the MMTV promoter and individual decay curves. (a) Schematic and decay curve for GR dissociation from site 1. Dotted line represents best fit to single exponential decay model. Solid line represents best fit to biphasic decay model. Fits covering the first 100 s of dissociation are shown in inset. b) Schematic and decay curve for GR dissociation from site 2. Symbols, lines and inset are as described above. GR concentration was 0.8 μM.

A GR/ER chimera also displays rapid and biphasic dissociation

To eliminate the possibility that our results were an artifact of our GR preparations, we measured the dissociation kinetics of a GR chimera, in which the ligand-binding domain of GR is substituted with that of estrogen receptor-α (Figure 3.8a). The chimera exhibits numerous
functional differences compared to GR, including differences in dimerization energetics, transcriptional activity and cooperativity between non-adjacent binding sites\textsuperscript{33}. Yet as shown in Figure 3.8b, GR/ER dissociation from the TAT\textsubscript{3} response element is again best described by a two-phase exponential decay model. The resolved fast and slow phase rate constant are also similar to those of GR, with $k_{obs,1}$ and $k_{obs,2}$ corresponding to $0.7 \pm 0.2 \text{ s}^{-1}$ and $0.014 \pm 0.004 \text{ s}^{-1}$, respectively. Analysis of GR/ER dissociation from a subset of the HREs in Figure 3.1 revealed statistically identical results (not shown). Collectively, our results indicate that biphasic dissociation kinetics are intrinsic to GR and GR/ER. From a mechanistic perspective, the results also indicate that molecular basis of biphasic decay is not obviously linked to the GR hormone-binding domain.

\textbf{Figure 3.8.} Schematic and decay curve for dissociation of the GR/ER chimera from the TAT\textsubscript{3} sequence. (a) Schematic representation of GR/ER primary structure. Ligand-binding domain of GR is replaced with that of ER-\alpha. Functional domains and regions are as indicated: NTR, N-terminal region; DBD, DNA-binding domain; LBD, ligand-binding domain. (b) Decay curve for GR/ER–TAT\textsubscript{3} dissociation. Open circles represent time-dependent GR/ER occupancy. Dotted line represents best fit to single-exponential model; solid line represents fit to biphasic decay model. GR/ER concentration was 0.7 μM.
Discussion

GR-DNA interactions occur on similar timescales both in vitro and in vivo

We report here that GR interactions with DNA are transient in vitro – receptor residence times last only seconds to tens of seconds. This seems to be a general result since we observe rapid exchange for full-length GR and a GR chimera, at an array of individual response elements, and at the multisite MMTV promoter. Recent surface plasmon resonance studies of GR DBD interactions with response elements are qualitatively consistent with this result, reporting apparent half-lives of ~20-50 seconds\textsuperscript{25}. Our findings therefore demonstrate that GR exchange with DNA occurs on a similar timescale in both the test-tube and in live cells. Although this convergence does not necessarily mean that the receptor uses identical DNA-binding mechanisms in both environments, it does suggest that additional biophysical studies of GR-DNA interactions in vitro should prove useful in interpreting receptor-chromatin behavior in vivo.

In contrast to the present results, early investigations found that GR-DNA residence times lasted minutes to hours\textsuperscript{60,61}. Although the basis of this discrepancy is unclear, we note that at the time of these studies, it was not yet possible to recombinantly express full-length GR. Instead, the receptor was extracted from tissue samples and partially purified in relatively low amounts. Although recombinant expression of the isolated DBD had been achieved, the extent of activity for the purified domain was not always clear. It is thus conceivable that in these early studies, contaminating proteins and/or structural heterogeneity contributed to long residence times.
Association kinetics control GR-DNA binding affinity

We previously found that for the six response elements shown in Figure 3.1a, GR binding affinity spanned a 80-fold range. Noting that the off-rates for these sequences are essentially identical, this implies that receptor on-rate controls binding affinity. This is somewhat surprising since studies of other transcription factors have shown that off-rate is the controlling factor. However, these studies were carried out using either “simple” repressors or fragments of more complex transcription factors. These examples may not be most appropriate for larger, multi-domain proteins such as GR. Better models may instead come from the field of immunology. Kinetic analyses of both antibody-epitope interactions and T cell receptor-peptide interactions have revealed that on-rate often controls binding affinity. Although the molecular basis for such control is not fully understood, it is thought that rate-limiting structural reorganization in one or both macromolecules must occur upon assembly. Interestingly, both GR and DNA are also known to undergo structural transitions upon binding. Whether these transitions are linked to on-rate control of affinity will require detailed studies of the time-dependence of GR-DNA association. We expect that these investigations (currently underway) will reveal complex and multiphasic kinetics, with association rate constants comparable to other transcription factors (∼1 x 10^8 M⁻¹ s⁻¹) that decrease as GR–DNA binding affinity weakens.

Biphasic GR-DNA dissociation suggests interconversion between states

Our final observation was that GR-DNA dissociation kinetics are biphasic in character. This indicates that GR does not dissociate in a single step, which would predict a mono-exponential decay curve. Instead, dissociation occurs in multiple steps, possibly via an intermediate. Our goal here is to establish a minimal kinetic model to describe this phenomenon. One possibility is that GR monomers sequentially dissociate from the half-sites within a response element. However,
this predicts that decay curves generated from each half-site should be different from each other, with one half-site showing rapid decay and the other showing slow. However, quantification of each half-site in the footprint shown in Figure 3.3a generated decay curves identical to each other and to the full-site curve shown in Figure 3.3b (not shown). Half-site analysis of the remaining response elements also revealed decay curves identical to those for the full sites. Although this concordance does not formally eliminate the possibility of sequential monomer dissociation (for example, monomers may randomly dissociate via a non-sequential pathway regardless of HRE sequence), we find no support for it.

A second possibility is that monomers randomly dissociate via a non-sequential pathway regardless of HRE sequence. Although we cannot formally eliminate this possibility, we think it unlikely for two reasons. First, we note that GR dimer binding to the DNA is associated with significant hypersensitivity adjacent to the binding site (Figure 3.3a); this is presumably due to receptor-induced DNA bending. If so, then hypersensitivity is reporting on the dimer-bound rather than monomer-bound ligation state. If monomers randomly and sequentially dissociate, we would therefore expect to observe a mono-exponential decay in hypersensitivity. Instead, quantification of the hypersensitive sites reveals biphasic decay identical to that seen in Figure 3.3b (not shown). Secondly, in contrast to our preferred model (described in more detail below), we can think of no obvious biological advantage to random monomer dissociation.

A third possibility is that in the absence of DNA, GR exists as two independent species existing at similar proportions. This might be due to the presence of a truncated or misfolded receptor population generated during purification. However, we would expect that these species would be manifested as heterogeneity in either our sedimentation velocity and/or equilibrium footprinting studies. Yet we detect a single species by sedimentation velocity and a
single binding transition by equilibrium footprinting. Noting that we also observe biphasic kinetics with the GR/ER chimera, this possibility also seems unlikely.

This led us to the model in Figure 3.9a, which postulates that the GR-DNA complex reversibly interconverts between two states, A and B, with state A dissociating from the DNA via rate constant $k_{\text{off},A}$. Interconversion between the states is controlled by rate constants, $k_{AB}$ and $k_{BA}$. Using the differential equations that describe this model (Eq 3.3-3.5) we globally fit eleven TAT$_3$ dissociation curves covering a ten-fold range of GR concentrations. The resultant fit is shown in Figure 3.9b, with the time-dependent probability of each DNA-bound state shown in Figure 3.9c. This latter plot predicts that at equilibrium ($t = 0$), states A and B exist in comparable proportions, with state A existing at just under 60% and state B slightly over 40%. With regard to kinetic parameters, the dissociation rate constant for state A, $k_{\text{off},A}$, is $0.4 \pm 0.1$ s$^{-1}$, corresponding to an average residence time of 2.5 s. More interesting are the values for the isomerization rate constants, $k_{AB}$ and $k_{BA}$. These were found to be $0.020 \pm 0.008$ s$^{-1}$ and $0.026 \pm 0.005$ s$^{-1}$, respectively. These translate to half-lives for each state of approximately 30 s.

Noting that GR interactions with coactivating proteins also occur on a second to tens of second timescale$^{93,94}$, we speculate that receptor interconversion on a similar timescale may have functional relevance. Similar to arguments made for T cell receptor-peptide interactions$^{95,96}$, it may be that the coactivating proteins and ligands that target GR are capable of distinguishing between the different DNA-bound states, leading to unique transcriptional responses – one possibility is that they reflect the productive versus unproductive transcriptional activation complexes observed in live cells$^{97}$. Alternatively, there may be synchronous cross-talk between receptor interconversion and cyclical chromatin remodeling events$^{98}$. Although further work will be necessary to confirm this thinking, we believe the model in Figure 3.9a represents the minimal complexity necessary to describe the data. Although other
models were consistent with our decay curves, they either required additional kinetic parameters or additional phases, or suggested no biological relevance, making them unwarranted.

\[\text{Figure 3.9.} \text{ Global analysis of GR-TAT}_3 \text{ decay curves using a minimal kinetic model. (a) DNA-bound GR interconverts between states A and B via rate constants } k_{AB} \text{ and } k_{BA}. \text{ GR dissociates only from state A via dissociation rate constant } k_{off,A}. \text{ Grey circles represent GR monomers and adjoined black squares represent response element. (b) Global analysis of eleven decay curves covering a ten-fold range of GR concentrations. Open circles represent time-dependent GR occupancy. Solid line represents best fit by numerical integration of the ordinary differential equations that describe model (see Materials and Methods). Inset shows fit to first 100 s of dissociation. (c) Probability of the two GR-DNA microstates in first 20 s. Dashed line represents probability of state A; dotted line represents the probability of state B. Solid line represents the sum of states A and B, and is equivalent to fit line in (a).}\]
Materials and Methods

Expression and purification of full-length human GR

Detailed protocols for expressing and purifying full-length GR have been described previously\textsuperscript{32,62}. Briefly, the receptor is expressed as a hexahistidine-tagged protein in baculovirus-infected insect cells. The nuclear lysate is fractionated over Ni\textsuperscript{2+}-agarose resin, followed by Sephacryl S-300 size exclusion chromatography. Eluted GR is then concentrated using Q-Sepharose. A similar approach was used to express and purify a glucocorticoid-estrogen receptor chimera (GR/ER) used here as a control protein, in which the ligand-binding domain of GR is substituted with that of human estrogen receptor-\textalpha\textsuperscript{33}. Saturating levels of the ligands triamcinolone acetonide (for GR), or 17-\textbeta-estradiol (for GR/ER) were present throughout receptor expression, purification and storage. Receptor concentrations were determined using calculated extinction coefficients of 71,280 M\textsuperscript{-1} cm\textsuperscript{-1} for GR and 52,830 M\textsuperscript{-1} cm\textsuperscript{-1} for GR/ER\textsuperscript{81}.

Sedimentation velocity analysis

Sedimentation velocity of GR was carried out as described previously, using a Beckman Optima XL-A analytical ultracentrifuge\textsuperscript{32,62}. Buffer conditions were 20 mM Tris-HCl (pH 8.0 at 4°C), 100 mM NaCl, 1 mM CaCl\textsubscript{2}, 2.5 mM MgCl\textsubscript{2}, 1 mM DTT, and 10 \mu M TA. GR was sedimented at 1.5 \mu M, 50,000 rpm and 4°C. Data were collected at 230 nm. Sedimentation coefficient c(s) distributions were determined using the program, Sedfit\textsuperscript{67}.

Quantitative footprinting - equilibrium

Equilibrium footprints were carried out as described by Ackers and co-workers\textsuperscript{42} with minor modifications\textsuperscript{43}. Footprinting was carried out using an 1100 bp DNA fragment containing the TAT\textsubscript{3} imperfect palindrome (see Figure 3.1a). Buffer conditions were identical to those in the
sedimentation velocity studies, with the addition of 100 µg/mL BSA and 2 µg/mL salmon sperm DNA. Binding isotherms were generated as previously described\textsuperscript{42,43} using the program, ImageQuant (Molecular Dynamics).

**Quantitative footprinting - kinetic**

Kinetic footprints were carried out as described by Beckett and Brenowitz\textsuperscript{99,100} with minor modifications. As an example, the above 1100 bp radiolabeled DNA fragment containing the TAT\textsubscript{3} sequence was equilibrated at a single GR concentration and under buffer conditions identical to those described for equilibrium footprinting. Upon reaching equilibrium, a 45 µL aliquot was transferred to a tube containing 5 µL of unlabeled HRE-containing DNA at 1 to 100-fold molar excess over GR (referred hereafter as “trap”). The sample was then allowed to incubate for times ranging from 2 to 600 seconds. After the desired amount of time was reached, 45 µL of the sample was transferred to a tube containing 5 µL of 0.6 units/µL DNase and allowed to react for exactly 3 seconds. DNase digestion was stopped by adding 45 µL of 20 mM EDTA and the sample was processed as described previously\textsuperscript{43}. This approach, when carried out over a range of trap incubation times, generated a single kinetic decay curve. We then repeated this process over a range of initial GR concentrations, for five additional HREs, and the MMTV promoter (Figure 3.1). Trap concentrations covering a 1 to 100-fold molar excess over GR concentration generated identical dissociation kinetics, indicating that re-association of GR to the radiolabelled DNA was not occurring and therefore only GR-DNA dissociation was being monitored.
Resolution of observable rate constants

All kinetic decay curves were first fit to single-exponential model:

\[ \bar{Y} = A \ e^{-k_{obs}t} \]  \hspace{1cm} (3.1)

where \( \bar{Y} \) is the fractional saturation of the DNA at time \( t \), \( A \) is the amplitude and \( k_{obs} \) is the observable rate constant. Noting that all time courses also showed visual evidence of biphasic decay, they were also fit to a double-exponential model:

\[ \bar{Y} = A_1 \ e^{-k_{obs,1}t} + A_2 \ e^{-k_{obs,2}t} \]  \hspace{1cm} (3.2)

where \( A_1 \) and \( A_2 \) are the amplitudes of the fast and slow phases, and \( k_{obs,1} \) and \( k_{obs,2} \) are their respective rate constants. Time courses were analyzed via non-linear least squares using the program Scientist (Micromath, Inc.) An F-test at the 95% confidence level was used to determine whether the double-exponential fit was statistically improved over the single-exponential.

Molecular interpretation of biphasic kinetics

Biphasic kinetics are indicative of a multistep dissociation process. A minimal kinetic model that best describes the data posits that the GR-DNA complex exists as an equilibrium between two interconverting states, A and B, with only state A dissociating from the DNA. As shown in Figure 3.9a, this model is defined by dissociation rate constant \( k_{off,A} \) and interconversion rate constants \( k_{AB} \) and \( k_{BA} \). Using this model, we globally fit eleven GR decay curves for the TAT_3 sequence using the following ordinary differential equations:

\[ \frac{d[D_f]}{dt} = k_{off,A}[GR_2 D]_A \]  \hspace{1cm} (3.3)

\[ \frac{d[GR_2 D]_A}{dt} = -(k_{AB} + k_{off,A})[GR_2 D]_A + k_{BA}[GR_2 D]_B \]  \hspace{1cm} (3.4)
\[
\frac{d[GR_2 D_B]}{dt} = k_{AB}[GR_2 D]_A - k_{BA}[GR_2 D]_B
\]  

(3.5)

Where \([D_f]\) is the time-dependent concentration of free DNA, \([GR_2 D]_A\) is the time-dependent concentration of the GR-DNA dimer complex in state A, and \([GR_2 D]_B\) is the time-dependent concentration of the complex in state B. At the start of dissociation, we set \([D_f] = 0\), \([GR_2 D]_A = 1/[1+(k_{AB}/k_{BA})]\), and \([GR_2 D]_B = 1-[GR_2 D]_A\). The expressions for \([GR_2 D]_A\) and \([GR_2 D]_B\), now expressed as fractions relative to each other, ensure that both states are in equilibrium at \(t = 0\).

Finally, the experimental observable – the fractional saturation as a function of time, or \(\bar{Y}(t)\), is equal to the time-dependent probability of the two states:

\[
\bar{Y}(t) = \frac{[GR_2 D]_A + [GR_2 D]_B}{[D_f] + [GR_2 D]_A + [GR_2 D]_B}
\]

(3.6)

Because the kinetic experiments were carried out over a range of GR concentrations, resulting in a range of initial \(Y\)-bar values, all time-courses were treated as transition curves \((Y_{app})\) with upper (m) and lower (b) end points:

\[
\bar{Y}(t)_{app} = (m-b) \cdot \bar{Y}(t) + b
\]

(3.7)

For presentation purposes, all decay curves shown herein were normalized to \(Y\)-bar values ranging from 0 to 1. Differential equations were solved implicitly using Scientist (Micromath, Inc).
CHAPTER IV
CONCLUSIONS

This thesis set out to examine two facets of steroid receptor function. First, to contribute to a quantitative understanding of receptor-specific function, the thermodynamics of wild type AR and T877A self-assembly and DNA-binding at a two-site promoter were examined. In chapter II, I found that both AR constructs show no evidence of dimerization and bind to a model two-site promoter with substantial intersite cooperativity. Compared to the other receptors examined in the lab, these results demonstrate that the steroid receptors partition their DNA-binding energetics in parallel with their phylogenetic divergence. For example, the more distantly related subgroup 3A member ER-α displays strong dimerization energetics and weak cooperativity. On the contrary, the subgroup 3C members PR-A, PR-B, GR, and wild type AR and T877A display either weak or no evidence of dimerization and are capable of engaging in cooperative interactions. As the steroid receptor family evolved from a single ER-like ancestral protein to two distinct subgroups (3A and 3C), the receptors appear to have lost the ability to dimerize strongly, yet gained the ability to participate in cooperative interactions. These findings suggest a thermodynamic framework of steroid receptor-specific function.

Under a thermodynamic model of steroid receptor specificity, it is assumed that transcriptional activity is proportional to the probability of observing an occupied promoter. This stems from the idea that the higher the probability a receptor saturates a promoter, the higher the probability that the promoter-bound receptor will recruit co-activating proteins and the basal transcriptional machinery to initiate transcription. The resolved thermodynamic parameters of steroid receptor interactions can help explain specificity with a relatively small number of rules. One way the receptors may be able to accomplish this is through differences in
self-association energetics. Consider the observation that AR displays no evidence of self-association, suggesting that AR can bind promoters comprised of half sites. On the other hand, both PR isoforms self-associate with dissociation dimerization constants of approximately micromolar, suggesting that PR isoforms may have a preference for promoters that consist of both full response elements and half sites. At a given protein concentration, the total AR concentration will be made-up of more monomers than that for PR. Under competitive conditions, AR will out-compete promoter architectures with enriched amounts of half sites as PR is inhibited by the presence of dimers. Thus, dimerization energetics could be one means by which AR achieves specificity over some of the other steroid receptors.

Another way AR could achieve specific function is through cooperative interactions. At the model two-site promoter, AR was found to cooperatively bind the template; whereas PR-A assembled onto this particular promoter architecture with significantly lower cooperativity. Under competitive conditions, as both receptors are titrated with equal amounts of receptor, the promoter is expected to be occupied primarily by AR. This particular promoter architecture could represent a layout that favors induction by AR and disfavors efficient assembly and the resultant transcriptional initiation by PR-A. However, this alone cannot explain how specificity is conferred between AR, GR and PR-B. These three receptors assemble at the model two-site promoter with statistically similar cooperativity terms, giving none of the three a selective advantage for this particular promoter architecture. The model two-site promoter could represent an architectural motif of promoters regulated by all three receptors.

Naturally-occurring promoters are not structured with only one specific distance between response elements, like the model two-site promoter described above, but show differences in distance and phasing between response elements. One way the receptors could establish receptor-specific occupancies is if the cooperativity for each receptor changes as a
function of intersite distance and intersite phasing. For example, recent work in our laboratory
demonstrates that PR-B cooperatively binds promoters with closely spaced response elements
that are in phase with one another (both sites are on the same face of DNA)\textsuperscript{76}. On the other
hand, GR maintains strong cooperativity over greater distances than PR-B and shows a different
dependence on phasing. Given that AR is a closely related receptor to GR and PR-B, AR
cooperativity could be differentially influenced by spacing and phasing between the two binding
sites, giving rise to AR-specific occupancy.

In addition to modulating cooperativity through intersite distance and phasing,
cooperativity can also be modulated by other promoter architectures, in particular, promoters
with half sites. For example, PR-A has been shown to assemble the half-sites of the MMTV
promoter with substantial cooperativity. On the contrary, PR-A assembled at the model two-site
promoter with a minimal amount of cooperativity. Thus, it’s feasible that different receptors
could also display different types of cooperativity at promoters containing half sites. In
combination with the other ways to occupy a promoter (self-assembly energetics, cooperativity
between full response elements), an variety of ways are available for promoters to be
configured such that certain promoter architectures result in the occupancy of a specific
receptor.

Taken with previous work on the steroid receptors, the results that I obtained for AR
suggest that self-assembly and DNA-binding interactions hold the potential to explain receptor-
specific function on a fundamental level. Ultimately, we might be able to explain the
complexities of receptor specificity with a set of rules and states that are encoded in the
promoter architecture. Other mechanisms are expected to contribute to receptor specificity;
however, the fundamental interactions such as these could be a predominant driving force.
Continuing to study these interactions could lead to a model of receptor-specific function based on the fundamental interactions between the receptors and their promoters.

At the start of this project, I originally proposed that the T877A point mutation would display different self-assembly energetics and/or a different degree of intersite cooperativity when compared to wild type AR. As mentioned in Chapter II, this is not the case; wild type AR and T877A exhibit statistically similar energetics of dimerization and cooperativity. However, these studies were performed with the synthetic steroid metribolone (R1881), a strong agonist of both AR constructs. Wild type AR and T877A can also bind to a variety of other steroids. For example, both constructs are activated by the natural steroids testosterone and dihydrotestosterone. Additionally, given the expanded ligand-binding pocket due to the T877A point mutation, this AR construct is also activated by a number of other steroids (e.g. progesterone and estrogen) and currently prescribed prostate cancer drugs (e.g. hydroxyflutamide, a metabolite of the drug flutamide)\textsuperscript{102}.

As a future direction, it would be informative to test if AR ligands are coupled to promoter assembly energetics. The agonists listed above show a spectrum of transcriptional activities and it would be useful to see if ligand is coupled to promoter assembly energetics for the Wild type receptor and to also see if ligand differentially modulates the two interaction energetics for T877A. To carry this out, first, it will be necessary to express and purify the AR constructs from baculovirus-infected Sf9 insect cells in the presence of the ligand of interest. Second, like the work presented in Chapter II, the self-assembly energetics should be evaluated with sedimentation velocity and equilibrium. Given that the T877A point mutation occurs in the ligand-binding domain, the region of the receptors largely responsible for dimerization, and that there was no difference in dimerization between this mutant and wild type, it seems unlikely that dimerization is linked to ligand type. But, because dimerization is couple to DNA-binding,
self-association should still be tested as a function of ligand. Again, similar to the work presented in Chapter II, cooperative energetics at the two-site model promoter should be examined as a function of ligand type with quantitative DNase footprinting. DNA-binding studies on the GR and the GR chimera GR/ER suggest that the ligand-binding domain is involved in intersite cooperativity at the model two-site promoter. Thus, it’s feasible that ligand will influence androgen receptor cooperativity. Subtle changes in cooperativity can dramatically influence receptor occupancy levels and small ligand-dependent differences could result in significantly different promoter occupancy levels.

For the second facet of steroid receptor function, I re-evaluated the in vitro dynamics of promoter assembly by the GR. To carry this out, I measured the dissociation kinetics of the receptor from a panel of six individual response elements and the natural mouse mammary tumor virus (MMTV) promoter. In chapter III, I determined that GR and GR/ER dissociate from DNA on a timescale of seconds, the same timescale observed in living cells. My results suggest that GR dissociation kinetics on a timescale of seconds are primarily dictated by the receptor-promoter interactions themselves. Although mechanisms in living cells, such as active displacement, have been shown to influence receptor residence time, my in vitro results suggest that the interaction between GR and DNA is the primary source determining the dissociation rate. This suggests that higher-order mechanisms may not be an absolute necessity to invoke dissociation on a second timescale.

I also found that GR dissociation kinetics are biphasic in character. As opposed to a rigid body-like receptor dissociating from a response element, biphasic dissociation indicates that something more complex is happening to the GR-DNA complex. A minimal model that is consistent with our data proposes that the GR-DNA complex is at equilibrium between two
isomerizing states, A and B, where GR dissociates from the DNA via state A. These distinct states may represent two activity states of the receptor-DNA complex.

The presence of two distinct states of the GR-DNA complex may help to explain the observation that of all the DNA-bound GR, less than 10% engages in transcriptionally productive interactions. It is generally thought that transcriptional activation occurs through a set of specific and ordered sequence of steps. Progression through these steps depends on productive events from the stochastic association of factors. Progress towards transcriptional activation only occurs when the right factor is recruited to the promoter region at the right time. For example, in the case of GR, the ligand-bound receptor translocates from the cytoplasm to the nucleus and searches the genome for its binding sites. Once bound to a response element, GR then recruits chromatin-modifying proteins and the basal transcriptional machinery in a stochastic and sequential manner. For example, if co-activating protein is not recruited by GR during the time the receptor is bound to DNA, the potential for transcriptional activation is eliminated and the process starts over with GR searching for its binding sites. GR will repeat this cycle of search and dissociation until the appropriate sequence of factors is recruited to initiate transcription. Downstream factors such as the basal transcriptional machinery are recruited similarly in a stochastic manner.

One possible explanation of my kinetic data is that one state of the GR-DNA complex represents a transcriptionally unproductive conformation such that factors required for the next step do not productively interact with the complex. In the other DNA-bound state, the GR-DNA complex could be in a productive conformation such that it proficiently recruits the factor(s) that are required in the path to transcriptional activation. The existence of two functionally distinct states reduces the probability that a productive interaction will occur between the receptor and the next factor required. The probability of positive interactions is already limited
by the highly transient nature of the GR-DNA interactions and the existence of two distinct states further reduces the probability that GR will productively interact with co-activating proteins. Current models of GR transcriptional activation assume that GR behaves as a rigid-like body and that the stochastic nature of GR-DNA transiency is the primary reason for the low level of transcriptionally productive complexes. My results, on the other hand, suggest that GR exists in two distinct states. It’s appealing to speculate that these two states carry distinct functionalities that contribute to transcriptionally productive and unproductive interactions.

The existence of distinct functional states could have evolved from a requirement for GR to signal that it has located a specific binding site within a promoter. By undergoing an isomerization between two distinct states, GR may be signaling to the cell that it is capable of binding co-activating proteins. It is thought that, in general, transcription factors search for their specific binding sites through facilitated diffusion. In facilitated diffusion, transcription factors sample non-target DNA by sliding and ‘hopping’ on local segments of DNA or transferring between distal segments. During this search process, because GR has not landed on a specific site, it does not signal for the recruitment of co-activating proteins. This could reduce the chance that GR will inadvertently activate or modulate a part of the chromatin structure during the search process for its specific binding sites.

These in vitro kinetics results indicate that the lower-level interactions between receptor and DNA are a primary source of dissociation. Because my dissociation results are comparable to those made in the context of a living cell, kinetic mechanisms proposed in the test tube may offer insight into the molecular mechanisms occurring in a living cell. The differences in conditions between the test tube and the cell are vast. However, current live-cell imaging technologies are not yet able to report on molecular mechanisms of steroid receptor action within a living cell. In vitro experiments have been performed for decades to this end.
and continuing kinetics experiments in the test tube for GR will increase our understanding of these fundamental interactions of transcriptional activation by GR and other transcription factors.

As a future direction, it would be useful to obtain direct evidence that the GR-DNA complex exists at equilibrium between two isomerizing states. Biphasic dissociation observed in Chapter III implies the presence of two states, but is not direct evidence for their existence. To test for the existence of these two interconverting states, single molecule Förster (or fluorescence) resonance energy transfer (smFRET) experiments could be carried-out. In a common smFRET experiment, a component of a macromolecular complex is immobilized onto a quartz slide after being labeled with two fluorophores: a donor and an acceptor. Upon excitation with a specific wavelength, the donor fluorophore absorbs the incident light and then processes the absorbed energy in one of two ways. If the donor fluorophore is more than 100 Å from the acceptor, the donor will absorb the light and fluoresce at a donor-specific wavelength. Or, if the donor is within 20 to 100 Å of the acceptor, the donor will absorb the light, transfer the energy to the acceptor fluorophore, which then fluoresces at an acceptor-specific wavelength. Because the two fluorophores fluoresce at different wavelengths, both conformations can be monitored with total internal reflection (TIR) microscopy. Following the time trajectories of individual molecules allows one to determine FRET efficiencies of each molecule. After many individual molecules have been monitored, one can determine the populations of distinct conformations (i.e. states) of the macromolecular complex. By performing experiments with this methodology, which bypasses the ensemble- and time-averaged parameters measured in traditional biochemical assays, the distinct isomerizing states of the GR-DNA complex could be monitored directly.
In the case of monitoring distinct conformations of DNA-bound GR, several items should be considered in the design of smFRET experiments: immobilization and fluorophore location. First, a component of the GR-DNA complex will need to be immobilized onto a quartz slide to allow for the monitoring of single molecules by TIR microscopy. One straightforward way to accomplish this is to immobilize an individual response element DNA sequence by biotin-streptavidin linkage.

The second consideration surrounds the labeling of GR with donor and acceptor fluorophores, a non-trivial task. Because the structure of full-length GR has not been resolved, optimizing the design of fluorophore locations will need to occur. Without a detailed structure available, it will not be possible to predict which constructs will fall within a distance that results in high FRET efficiency. To overcome this, one could examine a small panel of GR constructs. First the N-terminus could be labeled with the donor fluorophore via the hexahistidine tag\textsuperscript{103}, introducing a genetically encoded aldehyde\textsuperscript{104}, or using a sortase-mediated reaction\textsuperscript{105}. Next, a panel of single point mutations of unnatural amino acids with site-directed mutations could be introduced to position the acceptor fluorophore in different positions throughout the protein\textsuperscript{106,107}.

One GR construct could consist of an acceptor fluorophore on the N-terminal side of the DNA-binding domain (DBD). The N-terminal region (NTR) of GR is thought to be largely unstructured. However, studies by Thompson and co-workers indicate that upon interacting with DNA, the N-terminal region takes on additional secondary and tertiary structure. These studies were performed in bulk solution and it was not possible to assess different populations of GR conformation states. If the increased secondary and tertiary structure represents one of the two states proposed in the minimal kinetic model, DNA-induced folding of the largely
unstructured NTR may bring the N-terminal fluorophore (donor) and the DBD fluorophore (acceptor) into close proximity, resulting in increased FRET efficiency for one of the states.

A second GR construct could consist of an acceptor fluorophore on the exterior surface of the ligand binding domain. The crystal structure of the ligand-binding domain has been resolved and one could choose a variety of locations on the exterior surface of the subdomain to place the acceptor fluorophore. Intramolecular interactions between the N-terminal region and the C-terminal LBD have been documented for other steroid receptors. If one of the states of the GR-DNA complex displays interactions between the N-terminus and the C-terminal LBD, the fluorophores would be brought closer together, resulting in an increase in FRET efficiency. As described here, carrying-out smFRET experiments will require substantial optimization but could be a useful tool for investigating the distinct states of the GR-DNA complex. In addition to this, one may also be equipped to study how an intrinsically disordered region of a human transcription factor confers its function. Such experiments will also lend insight into how this disordered region behaves in the context of its well-structured neighboring domains.

In addition to investigating distinct GR-DNA states with smFRET, it would also be insightful to examine how GR dissociation is influenced by the presence of co-activating proteins. Recall that the minimal kinetic model proposed in Chapter III proposes that DNA-bound GR is involved in an isomerization reaction between two states, A and B. One state could effectively recruit a co-activating protein and the other state could be rendered incapable to recruit a co-activating protein. If this is the case, an excess of co-activating protein will stabilize the productive transcriptional activation complex, the complex capable of recruiting co-activating proteins. Because the population of the two states will be dominated by the transcriptionally competent state, dissociation is expected to appear as a one-phase exponential decay curve instead of a two-phase decay curve observed in Chapter III. To test this, one would
first need to express and purify a co-activating protein that interacts directly with GR. Steroid receptor co-activator 2 (SRC-2) has been shown to interact with GR and has been expressed and purified previously in our laboratory, making it a straightforward protein to isolate. Other possible proteins that have been shown to interact directly with GR are SRC-1 and TATA binding protein. Second, to test if purified SRC-2 (or another protein) interacts directly with GR, sedimentation velocity experiments should be performed as a function of co-activator concentration. Finally, to test if SRC-2 influences GR dissociation, time-resolved DNase footprinting should be performed with an excess of SRC-2 concentration over GR concentration. Although these dissociation experiments may not provide the structural detail as nuclear magnetic resonance or smFRET, the experiments outlined above would provide support for the claim that GR is bound to DNA in two isomerizing states, each with different functional abilities.

Despite being disparate chapters, a common theme exists between the two. My studies on the androgen receptor widened the picture of how steroid receptor-promoter interactions might play a role in receptor specificity. By studying full-length androgen receptors under a set of standard buffer conditions, I have contributed to a rigorous comparative analysis of all of the receptor family members in hopes of elucidating mechanisms that dictate specificity for the family. Results presented in Chapter II suggest that low-level interactions between the steroid receptors and their DNA-binding sites are evolutionarily conserved and hold the potential to explain receptor specificity. In other words, interactions as fundamental as receptor-receptor and receptor-DNA interactions can help explain the complexities of receptor-specific function. Although higher-level interactions may play a role in specificity, this increased complexity may not need to be required to explain specificity. My kinetic studies on the glucocorticoid receptor follow a similar trend. Studies on GR suggest that transcriptionally productive and unproductive interactions may be guided by GR-DNA interactions. Other processes within a cell will influence
transcriptional activation; however, the interactions between the receptor and its DNA-binding sites may be a primary mechanism of transcriptionally productive and unproductive interactions. Taken together, the two studies suggest that the fundamental interactions between the steroid receptors and their DNA-binding sites dictate specificity and productive versus unproductive transcriptional complexes; two phenomena whereby more complex mechanisms may have been given too much emphasis. Continuing to study steroid receptor-DNA interactions may eventually demonstrate that these fundamental interactions could be playing major roles in other areas such as development, tissue-specific expression, and steroid receptor-related diseases.
REFERENCES


In Chapter III, I re-evaluated the *in vitro* kinetics of GR dissociation from DNA by examining GR dissociation from a panel of six hormone response elements (HREs) and the natural mouse mammary tumor virus (MMTV) promoter with the full-length receptor. In addition to observing that GR dissociation kinetics occur on the second timescale, I also found that GR dissociation is biphasic in character. Biphasic dissociation led me to propose a minimal kinetic model of GR dissociation from DNA. This minimal model proposes that the GR-DNA complex exists under an equilibrium of two isomerizing states, A and B, where GR dissociates from DNA through state A only.

However, there are two concerns associated with this work. The first concern is that the reported dissociation kinetics were measured by-hand, which limits the speed by which the kinetics could be monitored. Despite using a rapid way to collect kinetics data by hand (instead of with automated equipment), it is not fast enough to capture a significant amount of data during the fast phase of the dissociation process. For example, a timepoint was obtained at 0 seconds, which corresponds to a fractional saturation of 1.0; however, the next timepoint at 2 seconds corresponds to fractional saturation values no more than 0.6. Thus, information describing the initial loss of DNA occupancy (the first 0.4 fractional saturation units) was lost due to the constraint of performing the experiments by hand. This lack of information during the initial period of GR dissociation could negatively impact the claim that GR dissociates from DNA in a biphasic manner, which is the basis for the minimal kinetic model of GR dissociation.
There are two potential reasons for the gap between the 0 second and 2 second
timepoints. First, the sudden drop in the fractional saturation could be due to the unlabeled
competitor DNA used to sequester unbound receptor. In Chapter III, dissociation experiments
were performed as a function of unlabeled competitor DNA concentration. The results indicated
that the dissociation kinetics were independent of trap concentration. Thus, biphasic
dissociation kinetics are not due to the unlabeled competitor DNA.

A second potential reason is that the mixing procedures used to generate timepoints at
2 seconds and longer give rise to biphasic dissociation kinetics. The act of mixing to obtain the
dissociation kinetics could have resulted in a substantial drop between the 0 second timepoint
and the two second timepoint. However, equilibrium footprinting results indicate that an issue
with mixing is not likely. As presented in Figure 3.2, equilibrium experiments were performed at
the TAT$_3$ response element using previously published procedures where rapid mixing was not
present. In addition to these experiments, equilibrium footprinting experiments were also
performed with a similar procedure as the kinetics experiments. For these equilibrium
footprints, I mixed the solutions of GR and radiolabeled DNA with the same strategy as the
kinetic footprints (i.e. 3-second exposure to buffer and a 3-second exposure to DNase). The only
difference between this second set of equilibrium footprints and the kinetic footprints is the
absence or presence of unlabeled competitor DNA, which was shown to not influence
dissociation (see above). The reported thermodynamics between the two equilibrium footprint
methods were statistically similar and the fractional saturation values across all receptor
centrations were comparable for both equilibrium footprint procedures. Additionally,
fractional saturation values are comparable to those at the 0 second timepoint of dissociation
progress curves for all receptor concentrations used. If mixing caused a sudden drop in DNA
occupancy, different fractional saturation values would have been observed among these
different experimental procedures. Thus, mixing should not influence the longer timepoints obtained in the kinetic experiments. The comparable fractional saturation values between different experimental procedures bolster the argument that GR-DNA dissociation kinetics are biphasic in character and helps to justify the proposed minimal kinetic model of GR dissociation.

The second concern associated with Chapter III involves the direct fitting of 11 progress curves of GR dissociating from the TAT₃ response element to the set of differential equations that describe the minimal kinetic model (Fig. 3.9; Eqns. 3.3 thru 3.7). The minimal kinetic model includes three kinetic parameters (k_AB, k_BA, and k_off,A) yet GR dissociation only displays two phases. Thus, is the data sufficient to resolve a unique set of parameters describing the minimal kinetic model?

To address this question, I generated error surfaces for each of the three kinetic parameters (k_AB, k_BA, and k_off,A). To resolve the parameters listed in Chapter III, all three of the parameters were floated during the global analysis of the 11 TAT₃ dissociation curves. Error surfaces were generated here by assuming a known value for one parameter (i.e. fixing k_AB at a given value) and resolving the remaining two parameters (i.e. k_BA and k_off,A) by a least squares analysis and reporting a standard deviation of the fit. The fitting process was repeated over a range of assumed values for each kinetic parameter. From these global fits, the error surfaces were constructed by plotting the standard deviation of fit as a function of the assumed parameter values (Figure A.1).

Visual inspection of Figure A.1 reveals that each error surface displays a minimum at the corresponding values that were resolved with the minimal kinetic model described in Chapter III. For the forward isomerization rate constant, k_AB, the error surface minimum corresponds to 0.020 s⁻¹, the same value presented in Chapter III. Attempts were made to increase the breadth of the assumed values below 0.007 s⁻¹ and above 1.4 s⁻¹. However, using a broad range of initial
guesses for the other two kinetic parameters (k_{BA} and k_{off,A}), the program Scientist (Micromath) could not establish a fit. Similarly, for the reverse isomerization rate constant, k_{BA}, the error surface minimum corresponds to 0.026 s\(^{-1}\), the same value presented in Chapter III. Also, the fitting program could not establish a fit for assumed k_{BA} values below 0.003 s\(^{-1}\) or above 0.14 s\(^{-1}\). Finally, the error surface associated with k_{off,A} also showed a minimum at 0.4 s\(^{-1}\), the same value in Chapter III. Although, a minimum exists, it is shallower than the other two surfaces. The fitting program could not establish a fit for assumed values below 0.12 s\(^{-1}\), but the program was able to establish fits for an assumed value up to 1,000 s\(^{-1}\). Increasing the parameter space up to 1,000 s\(^{-1}\) results in a flat error surface (standard deviations are identical to the standard deviations assuming k_{off,A} equals 10s\(^{-1}\)). For clarity, the x-axis of Figure A.1.c was not extended to this large k_{off,A} value.

This supplemental analysis of the TAT\(_3\) dissociation data provides support that a unique set of kinetic parameters have been resolved for the minimal kinetic model presented in Chapter III. It should be noted that the entire parameter space is infinite and cannot be examined to completion. This however should not take away from the point that this analysis provides additional confidence that the analysis presented in Chapter III resolves a unique set of parameters describing a minimal kinetic model of GR dissociation.

Taken together, equilibrium footprint experiments and the parameter error surfaces bolster the argument that GR dissociation is biphasic in character and is well described by a minimal kinetic model containing a unique set of parameters. Future kinetic experiments may reveal details of GR dissociation from its response elements, but the data presented here indicate a more complex situation and represent the minimal amount of complexity required to adequately describe the data.
Figure A.1. Error surfaces corresponding to the three parameters of the minimal kinetic model. Standard deviation of the fit in fractional saturation units is plotted as a function of assumed values of (a) $k_{AB}$, (b) $k_{BA}$, and (c) $k_{off,A}$. Open circles represent the standard deviation of fit for a given assumed parameter value. For clarity, the assumed kinetic parameters are displayed in semilog units.