Title
Structural srudies of the human estrogen receptors and the Escherichia coli 90 kilodalton heat shock protein

Author
Shiau, Andrew Kwan-Nan

Publication Date
1999

Peer reviewed|Thesis/dissertation

Structural Studies of the Human Estrogen Receptors and the Escherichia coli 90 Kilodalton Heat Shock Protein

by

Andrew Kwan-Nan Shiau

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Biochemistry

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA SAN FRANCISCO

[redacted]

Date

University Librarian

Degree Conferred: .....

Acknowledgments

When I started my studies at UCSF in 1991, I naively assumed that my graduate school experience would center around science - the experiments I would perform, the courses I would take, the scientific knowledge I would gain. Now as I end them in 1999, I realize that my assumption was incorrect. For me, graduate school has been about the people with whom I shared the various aspects of the experience (scientific and otherwise). I know that I would have not survived the journey had it not been for the support of many people and I am grateful that I now have this opportunity to thank them.

My thesis advisor, Dave Agard, one of the most intelligent people I have ever met, has been a constant source of scientific insight and support. Over the many years, the long conversations we shared in his office, teasing apart data, analyzing the literature, building and tearing down models, truly shaped my scientific reasoning. I thank him for letting me pursue projects outside the mainstream of his lab and for his endless enthusiasm for science, which really kept me going during the rough times. I am also grateful to the other two members of my thesis committee, Robert Fletterick and Keith Yamamoto. I thank Robert for teaching me about the politics of science and enjoying the present. Keith has always been there to give me sage advice at critical points in my scientific career - when I was preparing for orals, during my thesis committee meetings, and when I was deciding what to do after graduate school.

Geof Greene and his lab at the University of Chicago have been fantastic collaborators. I thank Geof and members of his lab (Lin Cheng, Dani Barstad,

and Paula Loria) for teaching me about the estrogen receptors and for their persistence in protein purification and crystal growing.

Various members of the Agard lab, past and present, deserve special mention. I thank David Baker, my first baymate in the lab, for his scientific acumen and his infectious curiosity. Julie Sohl was always there to listen to my scientific and personal woes (she should have charged me by the hour!). I am also grateful to Nick Sauter, Sheila Jaswal, and Yoko Haga for their invaluable assistance at one time or another. I owe Alan Derman and Ted Mau, my two closest friends in the lab, my undying thanks simply for putting up with me and my emotional outbursts over the years. In addition, I thank Alan for his excellent editorial skills and Ted for his unbelievably selfless help during my various trips to the synchrotron.

My colleagues in the Macromolecular Structure Group have made the 10th floor of the Medical Sciences Building both an interesting and fun place to work. Special thanks go to the lunch club (Karen Han, Sarah Gillmor, and Richard Wagner) for providing well-needed nourishment and companionship and to Paul Foster, Bob Keenan, Sherry Laporte, Kinkead Reiling, and Manish Butte for their experimental help and advice and our many late-night conversations.

I was truly fortunate to find many friends among my classmates and the other graduate students in the Biochemistry department. I thank Paul Peluso for being good in bed; Kevin Lustig for our many adventures in the velour Rabbit; and Hernan Espinoza, Kathleen Liu, and Jim Wilhelm for many nights of X-Files and video games. I thank Arshad Desai, Kayvan Roayaie, and Ramon Tabtiang for the late night problem set sessions, the pasta and card games at 4:00 A.M., and the conversations lasting until dawn we shared our first year. My deepest

gratitude also goes to Kayvan for being the best roommate I've ever had and for helping me deal with the aftermath of my bike accident and various surgeries.

I probably wouldn't have entered graduate school had it not been for my parents' emphasis on education (especially math and science) and I thank them for their unflagging support and understanding over these last eight years.

Finally, I thank Heather Deacon, the love of my life, for being my artist, my travelling companion, my confidant, and my best friend.

Structural Studies of the Human Estrogen Receptors and the *Escherichia coli* 90 Kilodalton Heat Shock Protein

by

Andrew Kwan-Nan Shiau

Abstract

The estrogen receptors (ERs), ERα and ERβ, are members of the nuclear receptor (NR) superfamily of ligand-regulated transcription factors. In the absence of ligand, the C-terminal ligand-binding domains (LBDs) of the ERs interact with a large, dynamic, chaperone complex formed around the 90 kilodalton heat shock protein (hsp90). This complex promotes the ability of the LBDs of the ERs to bind ligands. Ligand binding causes the dissociation of the LBDs from this chaperone complex. Ligand-dependent activation of transcription by the ERs is mediated by interaction with coactivators. Receptor agonists favor coactivator binding and receptor antagonists block coactivator binding. This thesis describes X-ray crystallographic studies aimed at determining how the structures of the LBDs of ERα and ERβ, in the unliganded state and in various liganded states, influence the interactions of the ERs with other proteins, in particular hsp90 and p160 coactivators. The structure of full-length HtpG, the *Escherichia coli* hsp90, reveals that HtpG shares an even greater structural homology to the topoisomerase, DNA gyrase B, and the DNA repair enzyme, MutL, than that previously reported. This structural homology suggests a possible model of how ATP-driven conformational changes of hsp90 may facilitate the ligand binding process. The binding of agonists drives a conformational change in the LBDs of

both ERs involving the repositioning of the most C-terminal helix of each of the LBDs, helix 12. Members of the p160 coactivator family, such as GRIP1, recognize the LBDs of the ERs through a small signature sequence motif, LXXLL (where L is leucine and X is any amino acid) called the NR box. A peptide based on the sequence of the second NR box of GRIP1 forms an amphipathic α-helix which binds in a shallow hydrophobic groove on the surface of the ERα LBD formed by helices 3, 4, 5 and 12. Antagonists block the interaction with coactivators by stabilizing alternate conformations of helix 12 that either partially or completely occlude this coactivator recognition groove.

David A. Agard, Ph.D.

Advisor

Table of Contents

List of Figures

Chapter 3

List of Tables

Chapter 1

The Structural Basis of Estrogen Receptor/Coactivator Recognition and the Antagonism of This Interaction by Tamoxifen

Andrew K. Shiau*, Danielle Barstad†, Paula M. Loria†, Lin Cheng†, Peter J. Kushner‡, David A. Agard*, and Geoffrey L. Greene†

* The Howard Hughes Medical Institute
and the Department of Biochemistry and Biophysics
University of California at San Francisco
San Francisco, California 94143

† The Ben May Institute for Cancer Research
and the Department of Biochemistry and Molecular Biology
University of Chicago
Chicago, Illinois 60637

‡ The Metabolic Research Unit
University of California at San Francisco
San Francisco, California 94143

The following chapter was originally published in *Cell* (Shiau, A. K., Barstad, D., Loria, P. M., Cheng, L., Kushner, P. J., Agard, D. A., and Greene, G. L., (1998). *95*, 927-937) and is reproduced here from the original, with permission (Appendix 1).

Cell, Vol. 95, 927–937, December 23, 1998, Copyright ©1998 by Cell Press

The Structural Basis of Estrogen Receptor/Coactivator Recognition and the Antagonism of This Interaction by Tamoxifen

Andrew K. Shiau,* Danielle Barstad,† Paula M. Loria,† Lin Cheng,† Peter J. Kushner,‡ David A. Agard,*§ and Geoffrey L. Greene†§
*Howard Hughes Medical Institute and the Department of Biochemistry and Biophysics University of California at San Francisco San Francisco, California 94143-0448
†The Ben May Institute for Cancer Research and Department of Biochemistry and Molecular Biology University of Chicago Chicago, Illinois 60637
‡Metabolic Research Unit University of California San Francisco, California 94143

Summary

Ligand-dependent activation of transcription by nuclear receptors (NRs) is mediated by interactions with coactivators. Receptor agonists promote coactivator binding, and antagonists block coactivator binding. Here we report the crystal structure of the human estrogen receptor α (hERα) ligand-binding domain (LBD) bound to both the agonist diethylstilbestrol (DES) and a peptide derived from the NR box II region of the coactivator GRIP1 and the crystal structure of the hERα LBD bound to the selective antagonist 4-hydroxytamoxifen (OHT). In the DES-LBD-peptide complex, the peptide binds as a short α helix to a hydrophobic groove on the surface of the LBD. In the OHT-LBD complex, helix 12 occludes the coactivator recognition groove by mimicking the interactions of the NR box peptide with the LBD. These structures reveal the two distinct mechanisms by which structural features of OHT promote this "autoinhibitory" helix 12 conformation.

Introduction

Estrogens exert their physiological effects by binding to the estrogen receptors, which are members of the nuclear receptor (NR) superfamily of ligand-inducible transcription factors (Tsai and O'Malley, 1994; Beato et al., 1995). The estrogen receptor α (ERα) regulates the differentiation and maintenance of neural, skeletal, cardiovascular, and reproductive tissues (Korach, 1994; Smith et al., 1994). Compounds that modulate ERα transcriptional activity are currently being used to treat osteoporosis, cardiovascular disease, and breast cancer (Gradishar and Jordan, 1997; Jordan, 1998).

All ERα ligands bind exclusively to the C-terminal ligand-binding domain (LBD). The LBD recognizes a variety of compounds diverse in their size, shape, and chemical properties. Some of these ligands, including the endogenous estrogen 17β-estradiol (E_2) and the synthetic nonsteroidal estrogen diethylstilbestrol (DES), function as pure agonists, whereas others, such as ICI-164,384, function as pure antagonists. Synthetic ligands such as tamoxifen and raloxifene (RAL) belong to a growing class of molecules known as selective estrogen receptor modulators (SERMs), which function as antagonists in specific tissue and promoter contexts (Grese et al., 1997). The remarkable tissue-specific behavior of tamoxifen was recently demonstrated in the National Surgical Adjuvant Breast and Bowel Project–sponsored Breast Cancer Prevention Trial (Smigel, 1998). In the group of women at high risk for breast cancer who received tamoxifen treatment, there was an increased incidence of endometrial cancer but a reduced occurrence of certain bone fractures and a dramatic 45% reduction in breast cancer incidence. The rational design of new SERMs and the optimization of existing ones require an understanding of the effects of different ligand chemistries and structures upon ERα transcriptional activity.

Transcriptional activation by ERα is mediated by at least two separate activation functions (AFs), AF-1 in the N terminus, and AF-2 in the LBD. The activity of AF-1 is regulated by growth factors acting through the MAP kinase pathway (Kato et al., 1995), while AF-2 activity is responsive to ligand binding (Kumar et al., 1987). The binding of agonists triggers AF-2 activity, whereas the binding of antagonists does not (Berry et al., 1990).

Recent structural studies suggest that ligands regulate AF-2 activity by directly affecting the structure of the LBD. Comparison of the structure of the unliganded human retinoid X receptor α LBD (Bourguet et al., 1995) with the structures of the agonist-bound LBDs of the human retinoic acid receptor γ (RARγ) (Renaud et al., 1995) and other NRs suggests that an agonist-induced conformational change involving the repositioning of helix 12, the most C-terminal helix of the LBD, is essential for AF-2 activity (Moras and Gronemeyer, 1998). Because certain point mutations in helices 3, 5, and 12 abolish AF-2 activity but have no effect on ligand or DNA binding, these regions of the LBD have been predicted to form part of a recognition surface, created in the presence of agonist, for molecules that link the receptor to the general transcriptional machinery (Danielian et al., 1992; Wrenn and Katzenellenbogen, 1993; Henttu et al., 1997; Feng et al., 1998).

The structures of the LBD complexed with E_2 and RAL show that, although both ligands bind at the same site within the core of the LBD (Brzozowski et al., 1997), each of these ligands induces a different conformation of helix 12. Whereas helix 12 in the E_2-LBD complex packs against helices 3, 5/6, and 11 in a conformation that has been observed for the corresponding helix in other agonist-bound NR LBD structures, helix 12 in the RAL-LBD complex is bound in a hydrophobic groove composed of residues from helices 3 and 5. This alternative orientation of helix 12 partially buries residues in the groove that are necessary for AF-2 activity, suggesting that RAL and possibly other antagonists block AF-2 function by disrupting the topography of the AF-2 surface.

§To whom correspondence should be addressed (e-mail: agard@msg.ucsf.edu or ggreene@huggins.bsd.uchicago.edu).

Several proteins, including SRC-1/N-CoA1 (Onate et al., 1995; Kamei et al., 1996), GRIP1/TIF2/N-CoA2 (Hong et al., 1996; Voegel et al., 1996; Torchia et al., 1997;), p/CIP/RAC3/ACTR/AIB1 (Anzick et al., 1997; Chen et al., 1997; Li et al., 1997; Torchia et al., 1997), and CBP/p300 (Hanstein et al., 1996), associate in a ligand-dependent manner with the ERα. These proteins have been classified as transcriptional coactivators because they enhance ligand-dependent transcriptional activation by the ERα as well as by several other NRs (Horwitz et al., 1996; Glass et al., 1997). SRC-1 and GRIP1 bind to the agonist-bound LBDs of both the human thyroid receptor β (TRβ) and human ERα using the putative AF-2 interaction surface (Feng et al., 1998). Members of the p160 family of coactivators, such as SRC-1 and GRIP1, as well as other coactivators, recognize agonist-bound NR LBDs through a short signature sequence motif, LXXLL (where L is leucine and X is any amino acid), known as the NR box (Le Douarin et al., 1996; Heery et al., 1997; Torchia et al., 1997; Ding et al., 1998). Mutagenesis studies indicate that the affinity of coactivators for NR LBDs is determined principally, if not exclusively, by these NR boxes (Le Douarin et al., 1996; Heery et al., 1997; Torchia et al., 1997; Ding et al., 1998).

The structural mechanisms by which binding of different ligands to ERα influences coactivator recruitment remain unclear. We have chosen to examine the structural and functional effects on the LBD of the binding of two chemically related compounds, the agonist, DES, and the selective antagonist, 4-hydroxytamoxifen (OHT), the active metabolite of tamoxifen (Grainger and Metcalfe, 1996). Here we report the 2 Å resolution structure of the ERα LBD bound to both DES and a peptide with the sequence of the second NR box (NR box II) from the p160 coactivator GRIP1, and the 1.9 Å X-ray crystal structure of the human ERα LBD bound to OHT. In the DES complex, the NR box peptide is bound in an α-helical conformation by a hydrophobic groove formed by residues from helices 3, 4, 5, and 12 and the turn between helices 3 and 4. In the OHT complex, rather than forming part of a functional AF-2 surface, helix 12 binds to and occludes the coactivator recognition site by mimicking the interactions formed by an NR box with the LBD. The two distinct mechanisms by which specific structural features of OHT direct this alternative conformation of helix 12 are discussed.

Results

Structure Determination

GRIP1, a mouse p160 coactivator, interacts both in vivo and in vitro with the ERα LBD bound to agonist (Ding et al., 1998), but not with the LBD bound to antagonist (Norris et al., 1998). Mutational studies of GRIP1 and its human homolog TIF2 suggest that of the three NR boxes from GRIP1, NR box II (residues 690 to 694) binds most tightly to the ERα LBD (Ding et al., 1998; Voegel et al., 1998). Competition assays indicate that a 13-residue peptide, NH_2-KHKILHRLLQDSS-CO_2H (residues 686 to 698 from GRIP1), containing NR box II, binds specifically to the agonist-bound ERα LBD ($IC_{50} < 0.4$ μM; P. J. K., unpublished) and to other agonist-bound NR LBDs (Darimont et al., 1998 [see Note Added in Proof]; Ding et al., 1998). In the present study, an electrophoretic mobility shift assay was used to demonstrate that the NR box II peptide bound the ERα LBD in the presence of the agonist DES but not the antagonist OHT (data not shown). In combination, these observations suggest that the NR box II peptide is a valid model for studying the interaction between GRIP1 and the ERα LBD.

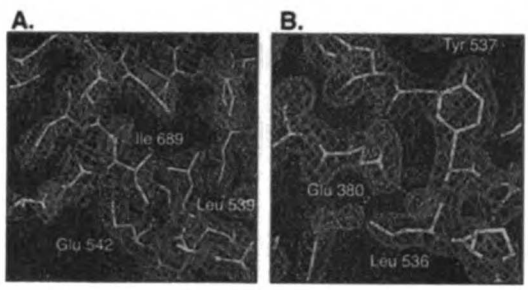


Figure 1. Views of the Electron Density of the DES-ERα LBD-GRIP1 NR Box II Peptide Complex and of the OHT-ERα LBD Complex

(A) A view of a $2F_o-F_c$ electron density map of the DES-LBD-peptide complex calculated at 2.03 Å resolution and contoured at 1.0 σ showing the GRIP1 NR box II interaction with the LBD. The peptide was omitted from the model prior to map calculation. Ile-689 from the peptide and two of the three receptor residues with which it interacts (Glu-542 and Leu-539) are labeled. Asp-538 has been omitted for clarity. The hydrogen bonds between the γ-carboxylate of Glu-542 and the amides of residues 689 and 690 of the peptide are depicted as dashed orange bonds.

(B) A view of a $2F_o-F_c$ electron density map of the OHT-LBD complex calculated at 1.90 Å resolution and contoured at 1.0 σ showing the N-terminal region of helix 12. The dashed orange bonds depict the water-mediated hydrogen bond network between the imidazole ring of His-377, the γ-carboxylate of Glu-380, and the amide of Tyr-537. The three labeled residues (Glu-380, Leu-536, and Tyr-537) interact with each other through van der Waals contacts and/or hydrogen bonds.

In order to characterize structurally the interaction between the GRIP1 NR box II peptide and the ERα LBD, recombinant human ERα LBD (residues 297–554) was crystallized bound to both DES and the peptide. The ERα LBD bound to OHT was also crystallized in order to determine the mechanism by which this antagonist blocks coactivator/ERα interaction. X-ray diffraction data from these crystals were measured, and the structures were determined by a combination of molecular replacement (using a modified version of the coordinates of the RARγ LBD [Renaud et al., 1995] as the search model) and aggressive density modification (see Experimental Procedures). The structure of the DES-ERα LBD-NR-box II peptide complex has been refined to a crystallographic R factor of 19.9% (R_{free} = 25.0%) using data to 2.03 Å resolution (Figure 1A and Table 1). The structure of the OHT-ERα LBD complex has been refined using data to 1.90 Å to a crystallographic R factor of 23.0% (R_{free} = 26.2%) (Figure 1B and Table 1).

Overall Structure of the DES-LBD-NR Box II Peptide Complex

The asymmetric unit of the DES-LBD-NR box II peptide complex crystals contains the same noncrystallographic dimer of LBDs that has been observed in the previously

Table 1. Summary of Crystallographic Statistics

Data Collection		
Ligand	DES	OHT
Space group	$P2_1$	$P6_522$
Resolution	2.03	1.90
Observations	104,189	269,253
Unique	30,265	23,064
Completeness (%)	98.4	99.1
R_{sym} (%)[a]	7.8	7.0
Average I/σI	9.8	16.1
Refinement		
No. of nonhydrogen atoms	4,180	2,069
R_{cryst} (%)[b]/R_{free} (%)	19.9/25.0	23.0/26.2
Bond rms deviation (Å)	0.005	0.006
Angle rms deviation (°)	1.05	1.04
Average B factor (Å²)	34.0	40.4

[a] $R_{sym} = \Sigma_i|I_i - \langle I_i\rangle|/\Sigma_i I_i$, where $\langle I_i\rangle$ is the average intensity over symmetry equivalents.
[b] $R_{cryst} = \Sigma||F_o| - |F_c||/\Sigma|F_o|$.

determined structures of the LBD bound to both E_2 and RAL (Brzozowski et al., 1997; Tanenbaum et al., 1998). The conformation of each LBD complexed with DES closely resembles that of the LBD bound to E_2 (Brzozowski et al., 1997); each monomer is a wedge-shaped molecule consisting of three layers of 11 to 12 helices and a single beta hairpin (Figure 2A). One NR box II peptide is bound to each LBD in a hydrophobic cleft composed of residues from helices 3, 4, 5, and 12 and the turn between 3 and 4 (Figures 2A and 3A). The density for both peptides in the asymmetric unit is continuous and unambiguous (Figure 1A). Residues 687 to 697 from peptide A and residues 686 to 696 from peptide B have been modeled; the remaining residues are disordered. Given that each peptide lies within a different environment within the crystal, it is striking that from residues Ile-689 to Gln-695 each peptide forms a two-turn, amphipathic α helix (Figures 2A and 3A). Flanking this region of common secondary structure, the peptides adopt dissimilar random coil conformations.

The NR Box II Peptide-LBD Interface

The binding of the NR box II peptide to the ERα LBD buries 1000 Å² of predominantly hydrophobic surface area from both molecules. The NR box II peptide-binding site is a shallow groove composed of residues Leu-354, Val-355, Ile-358, Ala-361, and Lys-362 from helix 3; Phe-367 and Val-368 from helix 4; Leu-372 from the turn between helices 3 and 4; Gln-375, Val-376, Leu-379, and Glu-380 from helix 5; and Asp-538, Leu-539, Glu-542, and Met-543 from helix 12 (Figure 3A). The floor and sides of this groove are completely nonpolar, but the ends of this groove are charged (Figure 3C).

The LBD interacts primarily with the hydrophobic face of the NR box II peptide α helix formed by the side chains of Ile-689 and the three LXXLL motif leucines (Leu-690, Leu-693, and Leu-694). The side chain of Leu-690 is deeply embedded within the groove and forms van der Waals contacts with the side chains of Ile-358, Val-376, Leu-379, Glu-380, and Met-543 (Figures 3A and 3C). The side chain of Leu-694 is similarly isolated within the groove and makes van der Waals contacts with the

Figure 2. Overall Structures of the DES-ERα LBD-GRIP1 NR Box II Peptide Complex and of the OHT-ERα LBD Complex

(A) Two orthogonal views of the DES-ERα LBD-NR box II peptide complex. The coactivator peptide and the LBD are shown as ribbon drawings. The peptide is colored gold, and helix 12 (residues 538–546) is colored magenta. Helices 3, 4, and 5 (labeled H3, H4, and H5, respectively) are colored blue. DES, colored green, is shown in space-filling representation.

(B) Two orthogonal views of the OHT-ERα LBD complex similar to those of the agonist complex in (A). The LBD is depicted as a ribbon drawing. As in (A), helix 12 (residues 536–544) is colored in magenta, and helices 3, 4, and 5 are colored blue. OHT, in red, is shown in space-filling representation.

side chains of Ile-358, Lys-362, Leu-372, Gln-375, Val-376, and Leu-379 (Figures 3A and 3C). In contrast, the side chains of both Ile-689 and the second NR box leucine, Leu-693, rest against the rim of the groove (Figures 3A and 3C). The side chain of Ile-689 lies in a shallow depression formed by the side chains of Asp-538, Leu-539, and Glu-542. The side chain of Leu-693 makes nonpolar contacts with the side chains of Ile-358 and Leu-539.

In addition to interacting with the hydrophobic face of the peptide helix, the LBD stabilizes the main chain conformation of the NR box peptide by forming capping interactions with both ends of the peptide helix. Glu-542 and Lys-362 are positioned at opposite ends of the peptide-binding site (Figure 3A). The γ-carboxylate of Glu-542 hydrogen bonds to the amides of the residues of N-terminal turn of the peptide helix (residues 688 and 689 of peptide A; residues 689 and 690 of peptide B) (Figure 1A). Similarly, the ε-amino group of Lys-362 hydrogen bonds to the carbonyls of the residues of the C-terminal turn of the peptide helix (residue 693 of peptide A; residues 693 and 694 of peptide B) (Figure 5). The side chain of Gln-375 also forms a water-mediated hydrogen bond to the carbonyl of residue 694.

To test the importance of the NR box peptide/LBD interface observed in the crystal, a series of site-directed mutations were introduced into the LBD. These mutations were designed either to perturb the nonpolar character of the floor of the binding groove (Ile-358→Arg,

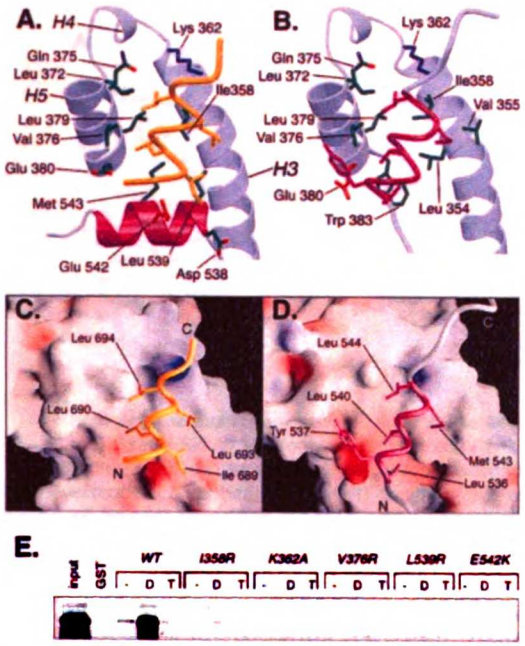


Figure 3. The NR Box II Peptide/DES-LBD Interface and the Helix 12/OHT-LBD Interface

(A) A close-up view of the coactivator peptide bound to the DES-LBD complex. The regions of the LBD that do not interact with the peptide have been omitted for clarity. Helices 3, 4, and 5 are labeled H3, H4, and H5, respectively. The side chains of the receptor residues that interact with the peptide are depicted. Except for Lys-362 (blue) and Glu-542 (red), the side chains are colored by atom type (carbon and sulfur atoms are colored green, oxygen atoms are colored red, and nitrogen atoms are colored blue). Helix 12 is colored magenta. The peptide, colored gold, is depicted as a Cα worm; only the side chains of Ile-689 and the three motif leucines (Leu-690, Leu-693, and Leu-694) are drawn.

(B) A close-up view of the OHT-LBD complex showing helix 12 bound to part of the coactivator-binding site. Only the side chains of residues that interact with helix 12 are drawn (with the exception of the side chain of His-373, which is omitted for clarity). Except for Lys-362 (blue) and Glu-380 (red), the side chains are colored by atom type (as specified in [A]). Residues 530–551 are depicted as a Cα worm; residues 536–544 are colored magenta. The side chains of Leu-536, Tyr-537, Leu-540, Met-543, and Leu-544 are shown.

(C) A molecular surface representation of the LBD bound to DES colored according to the local electrostatic potential (blue = positive; red = negative) as calculated in GRASP (Nicholls et al., 1991). The coactivator peptide is depicted as in (A) and the view is equivalent to that in (A). The side chains of Leu-690 and Leu-694 are bound in a hydrophobic groove, and those of Ile-689 and Leu-693 rest against the edge of this groove.

(D) A molecular surface representation of the LBD bound to OHT colored as in (C). Residues 530–551 are depicted as in (B) and the view is equivalent to that in (B). Whereas the side chains of Leu-540 and Leu-544 are embedded in the hydrophobic groove, that of Met-543 lies along the edge of this groove.

(E) ^{35}S-labeled GRIP1 was incubated with either immobilized glutathione S-transferase (GST), immobilized wild-type GST-hERα LBD (WT), or immobilized mutant GST-LBDs in the absence of ligand (−) or in the presence of DES (D) or OHT (T). Thirty picomoles of each of the GST-LBDs were immobilized, as described in the Experimental Procedures. The bound GRIP1 was visualized by fluorography after SDS-PAGE. The input lane represents the total amount of ^{35}S-GRIP1 included in each binding reaction. All of the mutations in the LBD disrupt agonist-dependent binding of GRIP1.

Val-376→Arg, and Leu-539→Arg) or to prevent the formation of the capping interactions (Lys-362→Ala and Glu-542→Lys) (Feng et al., 1998). Fusions of glutathione S-transferase (GST) to the wild-type and mutant LBDs were analyzed for their ability to bind ^{35}S-labeled GRIP1 in the absence of ligand or in the presence of DES or OHT. Only the wild-type GST-LBD was able to recognize the coactivator in the presence of DES (Figure 3E), confirming the importance of the observed capping and hydrophobic packing interactions.

Agonist Recognition

In its receptor complex, DES, like E_2 (Brzozowski et al., 1997), is completely encased within the narrower half of the LBD in a predominantly hydrophobic cavity composed of residues from helices 3, 6, 7, 8, 11, and 12 as well as the S1/S2 hairpin (Figures 2A and 4A).

The interaction of DES with ERα resembles that of E_2. One of the phenolic rings of DES lies in the same position as the E_2 A ring near helices 3 and 6. Like the aromatic ring of the E_2, the DES A ring (Figure 4A) is engaged by the side chains of Phe-404, Ala-350, Leu-387, and Leu-391 with its phenolic hydroxyl forming hydrogen bonds to the γ-carboxylate of Glu-353, to the guanidinium group of Arg-394, and to a structurally conserved water molecule. The A′ ring of DES (Figure 4A) is bound near helices 7, 8, and 11 adjacent to the location of the E_2 C and D rings. This ring forms van der Waals contacts not only with Gly-521 and Leu-525, like the D ring of E_2, but also with Met-343, Leu-346, and Met-421 (Figure 4A). Even though it is located 1.7 Å from the position of the D ring hydroxyl, the DES A′ ring phenolic hydroxyl is still able to hydrogen bond to the imidazole ring of His-524 (Figure 4A).

DES also forms contacts with the LBD that E_2 does not. There are unoccupied cavities adjacent to the α face of the B ring and the β face of the C ring of the E_2 (Brzozowski et al., 1997; Tanenbaum et al., 1998). The ethyl groups of DES, which project perpendicularly from the plane of the phenolic rings, fit snugly into these spaces. The resulting additional nonpolar contacts with the side chains of Ala-350, Leu-384, Phe-404, and Leu-428 (Figure 4A) may account for the higher affinity of DES for the receptor (Kuiper et al., 1997).

Except for Met-421 and Met-528 (both of which contact only DES) and Met-388 and Ile-424 (both of which contact only E_2), the ER is able to use the same residues to form all of the observed hydrogen bonds and van der Waals contacts with both of these distinctly shaped agonists (Figure 4A and Brzozowski et al., 1997; Tanenbaum et al., 1998). This remarkable adaptability is presumably the result of both the relatively large molecular volume of the binding pocket (~500 Å^3 in both complexes) and its apparent structural plasticity. In particular, at the DES A′ ring/steroid D ring end of the binding pocket, Met-343, Met-421, His-524, and Met-528 adopt different packing configurations in response to each ligand (data not shown).

Structure of the OHT-LBD Complex

The binding of OHT induces a conformation of the LBD that differs in both secondary and tertiary structural organization from that driven by DES binding. In the DES

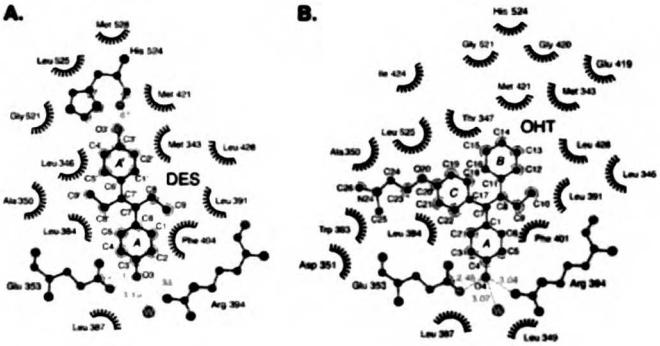


Figure 4. DES Interactions with the LBD (A) and OHT Interactions with the LBD (B)

Residues that interact with the ligands are drawn at approximately their true positions. The residues that form van der Waals contacts with ligand are depicted as labeled arcs with radial spokes that point towards the ligand atoms with which they interact. The residues that hydrogen bond to ligand are shown in ball-and-stick representation. Hydrogen bonds are represented as dashed cyan lines; the distance of each bond is given. The ligand rings and the individual ligand atoms are labeled.

complex, the main chain from residues 339 to 341, 421 to 423, and 527 to 530 forms parts of helices 3, 8, and 11, respectively. In contrast, these regions adopt an extended conformation in the OHT complex (Figures 2A, 2B, and 6A). In addition, the composition and orientation of helix 12 are different in the two structures. Helix 12 in the DES complex consists of residues 538 to 546, whereas helix 12 in the OHT complex consists of residues 536 to 544. Most dramatically, rather than covering the ligand-binding pocket as it does in the DES complex, helix 12 in the OHT complex occupies the part of the coactivator-binding groove formed by residues from helices 3, 4, and 5 and the turn connecting helices 3 and 4 (Figures 2A, 2B, and 3B). This alternative conformation of helix 12 appears to be similar to that observed in the RAL complex (Brzozowski et al., 1997).

Helix 12–LBD Interface

Except for the orientation of helix 12, the structure of the peptide-binding groove is almost identical in the DES-LBD-NR box II peptide, OHT-LBD, and E_2-LBD complexes (Figures 3A and 3B) (Brzozowski et al., 1997). We therefore refer to the region of this groove outside of helix 12 as the "static region" of the NR box–binding site (Feng et al., 1998). Helix 12 in the OHT complex and the NR box peptide helix in the DES complex interact with the static region of the coactivator recognition groove in strikingly similar ways.

Helix 12 mimics the hydrophobic interactions of the NR box peptide with the static region of the groove with a stretch of residues (residues 540 to 544) that resembles an NR box (LLEML instead of LXXLL). The side chains of Leu-540 and Met-543 lie in approximately the same locations as those of the first and second motif leucines (Leu-690 and Leu-693) in the peptide complex (Figure 5). Leu-540 is inserted into the groove and makes van der Waals contacts with Leu-354, Val-376, and Glu-380 (Figures 3B and 3D). Met-543 lies along the edge of the groove and forms van der Waals contacts with the side chains of Leu-354, Val-355, and Ile-358 (Figures 3B and 3D). The side chain position of Leu-544 almost exactly overlaps that of the third NR box leucine, Leu-694 (Figure 5). Deep within the groove, the Leu-544 side chain makes van der Waals contacts with the side chains of Ile-358, Lys-362, Leu-372, Gln-375, Val-376, and Leu-379 (Figures 3B and 3D).

Helix 12 in the OHT complex is also stabilized by N- and C-terminal capping interactions. Lys-362 interacts with the C-terminal turn of helix 12 much as it does with the equivalent turn of the peptide helix (Figures 3A and 3B). The Lys-362 side chain packs against the C-terminal turn of helix 12 with its ϵ-amino group hydrogen bonding to the carbonyls of residues 543 and 544 (Figure 5). Given that the capping interaction at the N-terminal turn of the coactivator helix is formed by a helix 12 residue (Glu-542), the N-terminal turn of helix 12 in the antagonist complex is stabilized by another residue, Glu-380 (Figures 3B and 3D). The Glu-380 γ-carboxylate forms van der Waals contacts with Tyr-537 and interacts with the amide of Tyr-537 through a series of water-mediated hydrogen bonds (Figure 1B).

In addition to forming these "NR box–like" interactions, helix 12 also forms van der Waals contacts with areas of the LBD outside of the coactivator recognition groove. The side chain of Leu-536 forms van der Waals contacts with Glu-380 and Trp-383, and that of Tyr-537 forms van der Waals contacts with His-373, Val-376,

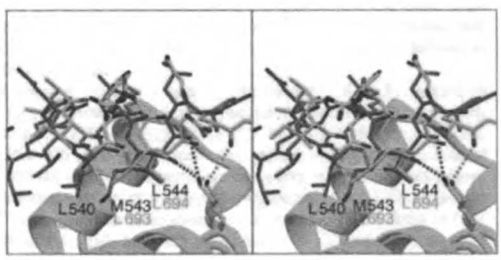


Figure 5. Comparison of Helix 12 from the OHT Complex and the NR Box II Peptide

The structures of the OHT-LBD complex and the DES-LBD-NR box II peptide complex were overlapped using the Cα coordinates of residues 306–526 of the LBD. Helix 12 from the DES-LBD-coactivator peptide complex is omitted for clarity. Residues 536–551 (helix 12 = residues 536–544) from the OHT-LBD complex are colored magenta, and the peptide is colored gold. For the OHT-LBD complex, the hydrogen bonds between the ϵ-amino group of Lys-362 and the backbone carbonyls of residues 543 and 544 of helix 12 are illustrated as dashed magenta bonds. For the DES-LBD-peptide complex, the hydrogen bonds between the ϵ-amino group of Lys-362 and the backbone carbonyls of residues 693 and 696 of the coactivator peptide are depicted as dashed orange bonds. Helix 12: L540 = Leu-540; M543 = Met-543; L544 = Leu-544. Peptide: L690 = Leu-690; L693 = Leu-693; L694 = Leu-694.

and Glu-380 (Figures 1B, 3B, and 3D). As a result of these contacts, helix 12 in the OHT complex buries more solvent-accessible surface area (~1200 Å^2) than the NR box peptide in the DES complex.

OHT Recognition

OHT is bound within the same pocket that recognizes DES, E_2, and RAL. The orientation of OHT within the binding pocket appears to be dictated by the positioning of two structural features of this ligand, the phenolic A ring and the bulky side chain (Figures 4B and 6C). The A ring of OHT is bound in approximately the same location as the A ring of DES near helices 3 and 6, with its phenolic hydroxyl hydrogen bonding to a structurally conserved water and to the side chains of Glu-353 and Arg-394 (Figure 4B). Like the bulky side chain of RAL, the side chain of OHT exits the binding pocket between helices 3 and 11 (Figures 2B and 4B). The OHT C ring (Figure 4B) forms van der Waals contacts with the side chains of Met-343, Leu-346, Thr-347, Ala-350, Trp-383, Leu-384, Leu-387, and Leu-525. The positioning of the flexible dimethylaminoethyl region of the side chain is stabilized by van der Waals contacts with Thr-347, Ala-350, and Trp-383 and by a salt bridge between the dimethylamino group of the side chain and the β-carboxylate of Asp-351, which lies 3.8 Å away (Figure 4B). The positions of the A ring and the side chain in the context of the rigid triphenylethylene framework of OHT requires that the ethylene group of OHT lie in an orientation nearly orthogonal to that of the ethylene group of DES (Figures 4A, 4B, and 6D). As a result, the B ring of OHT is driven more deeply into the binding pocket than the A′ ring of DES (Figures 6B and 6C).

This location of the OHT B ring apparently cannot be accommodated by the same mechanisms that allow the DES A′ ring/E_2 D ring end of the binding pocket to adapt to the different structural features of DES and E_2. Instead, the residues that contact the B ring (Met-343, Leu-346, Met-421, Ile-424, Gly-521, His-524, and Leu-525), most of which also interact with the A′ ring of DES, adopt conformations distinct from the ones they adopt in the DES structure (Figure 6D). In fact, the location of the B ring actually precludes the side chain of one residue, Met-421, from adopting the same conformation that it adopts in the DES structure (Figures 6B and 6C). As a consequence of these B ring–induced side chain conformations, many interresidue van der Waals contacts present in the DES complex are absent in the OHT complex. For example, whereas Met-421 packs against His-524 from helix 11 and against Met-343 from helix 3 in the agonist complexes, it is precluded by the location of the OHT B ring from interacting with either of these residues in the antagonist complex (Figure 6D).

The structural effects of the placement of the B ring are not limited to the residues that contact the B ring; the conformations of these residues force other residues throughout the binding pocket to, in turn, adopt alternative conformations. For instance, the conformation of Met-421 in the OHT complex prevents the side chains of Phe-404 and Phe-425 from occupying the positions they take in the DES complex (Figures 6B and 6C). The alternative conformations of the side chains of both the residues that directly contact the B ring and those that are indirectly affected by it force the main chain throughout the binding pocket to adopt a different conformation as well (Figure 6D).

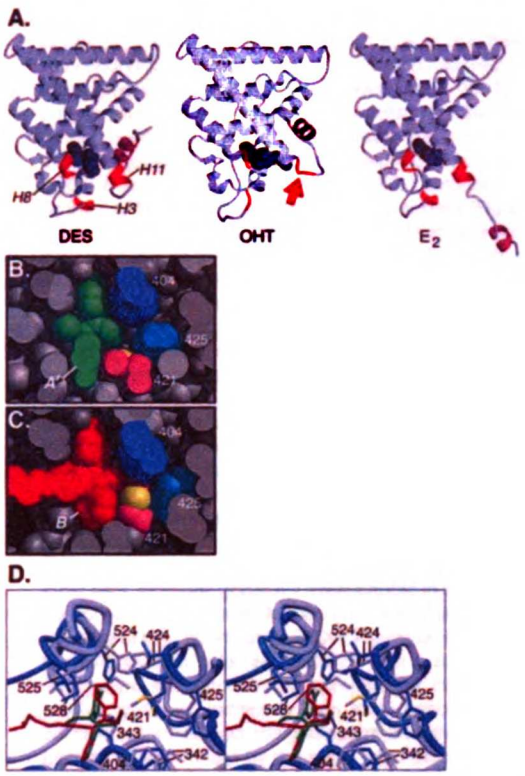


Figure 6. The Binding of Agonists and Antagonists Promote Different LBD Conformations

(A) Ribbon representations of the DES complex (without the coactivator peptide), the OHT complex, and the E_2 complex of Tanenbaum et al. (1998). The hormones are shown in space-filling representation. In each complex, helix 12 is colored magenta, and the main chain of residues 339 to 341, 421 to 423, and 527 to 530 is colored red. Helices 3, 8, and 11 (H3, H8, and H11, respectively) are labeled in the DES complex.

(B) A cross section of a space-filling model of the LBD bound to DES (green) showing the ligand completely embedded in the ligand-binding cavity. The A′ ring of DES (A′), Phe-404 (404), Met-421 (421), and Phe-425 (425) are labeled. The carbon atoms of side chain of Met-421 are colored magenta, and the sulfur atom is colored yellow.

(C) A cross section of a space-filling model of the LBD bound to OHT (red). The view is equivalent to that in (B). The B rings of OHT (B), Phe-404 (404), Met-421 (421), and Phe-425 (425) are labeled. The side chain of Met-421 is colored as in (B). The conformation of the B ring forces Met-421 to adopt a different conformation than the one it adopts in the DES complex (compare with [B]).

(D) The structures of the OHT complex and the DES complex were overlapped as in Figure 5. OHT is colored red, and DES is colored green. The A rings of both ligands point out of the page; the B ring of OHT and the A′ ring of DES point into the page. The LBD bound to OHT is colored blue, and the LBD bound to DES is colored light gray. The side chains of some of the residues whose conformations are dramatically different between the two complexes are drawn: Met-342 (342); Met-343 (343); Phe-404 (404); Met-421 (421); Ile-424 (424); Phe-425 (425); His-524 (524); Leu-525 (525); Met-528 (528). The sulfur atom of Met-421 is colored yellow in both structures.

Discussion

The AF-2 Surface and NR Box Recognition

The structure of the ERα LBD in complex with the GRIP1 NR box II peptide reveals that the LXXLL motif forms the core of a short amphipathic α helix that is recognized by a highly complementary groove on the surface of the receptor. In agreement with the conclusions of other mutational and structural studies (Brzozowski et al., 1997; Feng et al., 1998), we propose that this peptide-binding groove formed by residues from helices 3, 4, 5, and 12 and the turn between helices 3 and 4 is the AF-2 surface of ERα.

Of the eleven AF-2 residues whose side chains interact with the coactivator helix (Figure 3A), only four (Lys-362, Leu-379, Gln-375, and Glu-542) are highly conserved across the nuclear receptor family (Wurtz et al., 1996). The side chains of Gln-375 and Leu-379 are predominantly buried even in the absence of GRIP1 binding and appear to form integral parts of the architecture of the AF-2 surface. In contrast, the side chains of Lys-362 and Glu-542 are largely solvent exposed in the absence of coactivator and make both nonpolar contacts and the only direct receptor-mediated polar interactions with the coactivator helix. These two capping interaction residues are perfectly positioned at opposite ends of the AF-2 surface groove not only to stabilize the main chain conformation of the coactivator but also to function as a molecular caliper; the 15 Å distance between Lys-362 and Glu-542 is well suited to measure off the ~11 Å axial length of the short, two-turn coactivator α helix (Figure 3C). Similar receptor-mediated capping interactions have also been observed in a complex between the TRβ LBD and the NR box II peptide (Darimont et al., 1998). Mutation of either of these two capping interaction residues severely cripples coactivator binding by ERα as well as by TRβ (see Results and Henttu et al., 1997; Feng et al., 1998). Hence, the formation of helix-capping interactions may be a general feature of coactivator recognition by NRs.

The hydrophobic face of the NR box helix is formed by the side chains of the three motif leucines and the isoleucine preceding the motif (Ile-689). The functional importance of the conserved leucines in receptor binding has been demonstrated by numerous studies (Le Douarin et al., 1996; Heery et al., 1997; Torchia et al., 1997; Ding et al., 1998; Voegel et al., 1998). Structural and biochemical data in this study implicate Ile-689 as another key ERα-binding determinant. In the crystal, only the side chains of the motif leucines and Ile-689 extensively contact the LBD in both noncrystallographic symmetry-related peptides. Mutation of Ile-689 to alanine reduces the ability of the NR box II peptide to inhibit the binding of GRIP1 to ERα by ~30-fold in a competition assay (data not shown). Remarkably, the residue preceding the LXXLL motif differs for the three NR boxes of GRIP1/TIF2. The sequence variability at this position may explain the apparently different affinities of the NR boxes of TIF2 for the ERα (Voegel et al., 1998).

Helix 12 and the Regulation of AF-2 Activity

ERα AF-2 activity is blocked by antagonists such as OHT and RAL. The most striking feature of the structures of the OHT- and RAL-liganded ERα LBDs is that helix 12 is bound to the static region of the coactivator recognition groove (Figure 3B and Brzozowski et al., 1997). A comparison of these two structures with the structure of the coactivator/LBD complex reveals that in the antagonist complexes, the region of helix 12 with an NR box–like sequence (LXXML versus LXXLL) functions as an intramolecular mimic of the coactivator helix (Figure 5 and Brzozowski et al., 1997). Consistent with the proposals of others (Brzozowski et al., 1997; Darimont et al., 1998), this disposition of helix 12 directly affects the structure and function of the AF-2 surface in two ways. First, because helix 12 residues form an integral part of the AF-2 surface, the AF-2 surface is incomplete when helix 12 is in the antagonist-bound conformation. In particular, Leu-539, Glu-542, and Met-543 are incorrectly oriented for coactivator recognition. Second, residues from the static region of the AF-2 surface are bound to helix 12 and are prevented from interacting with coactivator (Figures 3A and 3B).

The sequence similarity of helix 12 of the ERα LBD to the LXXLL motif is not shared by all other NRs; the identities of the residues in this region of helix 12 in most NRs, although generally hydrophobic in character, do not as closely resemble the sequence of an NR box as those of ERα (Wurtz et al., 1996). However, it is possible that an intramolecular inhibitor with a suboptimal recognition sequence would compete for coactivator binding given its extremely high local concentration. Therefore, it will be necessary to determine if the antagonists of other NRs act by the same mechanism.

The Structural Basis of OHT Antagonism

The binding of OHT to ERα promotes a helix 12 conformation that inhibits binding of coactivator. OHT does not directly interact with any helix 12 residues (Figure 4B). Moreover, the structure of the LBD in the region of the AF-2 surface groove that interacts with helix 12 in the OHT complex is the same in the DES and E_2 complexes (Figures 3A, 3B, and 5) (Brzozowski et al., 1997). So how does OHT binding influence the position of helix 12?

Numerous studies have demonstrated the importance of the OHT side chain in receptor antagonism (Jordan and Gosden, 1982; Robertson et al., 1982). A comparison of the structures of the OHT and DES complexes reveals that the binding mode of the OHT side chain precludes the agonist-induced conformation of helix 12. The OHT side chain projects out of the ligand-binding pocket between helices 3 and 11 (Figures 2B, 6B, and 6C). As a result, the positioning of helix 12 over the ligand-binding pocket, as it is in the agonist-bound conformation, would bury the positively charged dimethylamino group of the OHT side chain within a hydrophobic cavity and produce steric clashes between the dimethylaminoethyl region of side chain and the side chain of Leu-540.

In functional terms, OHT is not, however, simply "an agonist with a side chain." OHT binding promotes a conformation of the LBD that is distinct from that stabilized by either DES or E_2 binding. These different conformations impose different restrictions on the positioning of helix 12.

Helices 3, 8, and 11 in the DES and E_2 complexes are between one to two turns longer than they are in the OHT complex (Figure 6A and Brzozowski et al., 1997). Helix 11 ends at Cys-530 in the DES and E_2 complexes and at Tyr-526 in the OHT complex. Helix 12 begins at Leu-536 in the OHT complex. This appears to be necessary; in the antagonist complex, Leu-536 forms a cooperative network of nonpolar contacts and hydrogen bonds with Glu-380 and Tyr-537 that stabilizes the N terminus of helix 12 (Figure 1B). Therefore, if helix 12 were to bind the static region of the AF-2 surface in the presence of agonist, the loop connecting helices 11 and 12 would be required to span ~17 Å over five residues. Although theoretically possible, this conformation would be highly strained and hence unlikely. In contrast, the longer loop connecting helices 11 and 12 in the OHT complex allows helix 12 to extend to the static region of the coactivator-binding groove.

In the DES and E_2 complexes, helix 12 and the loop connecting helices 11 and 12 pack against helices 3 and 11, whereas they do not in the OHT complex (Figures 2A and 2B and Brzozowski et al., 1997). Are the longer helices in the DES and E_2 complexes dependent upon the interactions helix 12 forms in the agonist-bound conformation? A recently described structure of the E_2-LBD complex suggests that they are not (Tanenbaum et al., 1998). In this structure, a crystal-packing artifact forces helix 12 to contact a symmetry-related molecule. Helix 12 is clearly not positioned over the ligand-binding pocket in this structure. Nevertheless, helices 3, 8, and 11 are longer than they are in the OHT complex (Figure 6A). Hence, the longer helices of the agonist complexes occur independently of the positioning of helix 12 over the ligand-binding pocket and are instead a direct result of agonist binding.

The secondary structure differences between the agonist complexes and the OHT complex arise from distinct arrangements of packing interactions induced by the different ligands. A cooperative network of van der Waals contacts, organized around DES or E_2, between various hydrophobic residues from helices 3, 7, 8, and 11 and the β hairpin appears to stabilize the longer helices in the agonist complexes (Figures 4A and 6D). The placement of the OHT B ring forces many of the ligand-binding pocket residues that surround it to adopt conformations that are dramatically different from those they adopt in either the DES or E_2 structures (see Results). As a result, many of the interresidue packing interactions present in the DES and E_2 structures are either absent or altered in the OHT structure (Figure 6D). These structural distortions apparently force the main chain from residues 339 to 341, 421 to 423, and 527 to 530 (which form parts of helices 3, 8, and 11, respectively, in the agonist structures) to adopt an extended conformation in the OHT structure (Figures 6A–6D).

Therefore, the binding of OHT has two distinct effects on the positioning of helix 12, each of which contributes to antagonism. Helix 12 is prevented from being positioned over the ligand-binding pocket by the OHT side chain. In addition, the alternative packing arrangement of ligand-binding pocket residues around OHT stabilizes a conformation of the LBD that permits helix 12 to reach the static region of the AF-2 surface and mimic bound coactivator.

These mechanisms do not appear to be specific to OHT. The side chain of RAL, like that of OHT, sterically hinders the agonist-bound conformation of helix 12 (Brzozowski et al., 1997). In addition, helix 11 appears to end at Met-528 in the RAL complex. This may result from the distortions in the binding pocket in the vicinity of His-524 directed by RAL binding (Brzozowski et al., 1997).

There is a great need for the improvement of existing therapies and the development of new ones for the prevention and treatment of breast cancer. While the tissue-selective antagonism of SERMs such as OHT and RAL is the result of numerous factors (Grainger and Metcalfe, 1996; Grese et al., 1997; Jordan, 1998), dissection of the mechanisms of action of these ligands requires a comprehensive understanding of how they act on the LBD and regulate its interactions with other cellular factors. Our studies have revealed, unexpectedly, that ligand-mediated structural perturbations in and around the ligand-binding pocket, and not simply side chain effects, contribute to receptor antagonism. Adjusting the balance between these two effects provides a novel strategy for the design of improved SERMs.

Experimental Procedures

Protein Expression and Purification

Human ERα LBD (residues 297–554) was expressed in BL21(DE3)-pLysS (harboring a plasmid provided by P. Sigler) as described previously (Seielstad et al., 1995). Bacterial lysates were applied to an estradiol-Sepharose column (Greene et al., 1980), and bound hERα LBD was carboxymethylated with 5 mM iodoacetic acid (Hegy et al., 1996). Protein was eluted with 3×10^{-5} M ligand in 30–100 ml of 50 mM Tris, 1 mM EDTA, 1 mM DTT, and 250 mM NaSCN (pH 8.5). The hERα LBD was further purified by ion exchange chromatography (Resource Q, Pharmacia). Protein samples were analyzed by SDS-PAGE, native PAGE, and electrospray ionization mass spectrometry.

GST-Pulldown Assays

A fusion between GST and amino acids 282–595 of hERα was constructed by subcloning the EcoRI fragment from pSG5 ERα-LBD (Lopez et al., submitted) into pGEX-3X (Pharmacia). Mutations were introduced into this construct using the QuikChange Kit (Stratagene) or by subcloning the appropriate fragments of mutant derivatives of pSG5-ER-HEGO (Tora et al., 1989; Feng et al., 1998). All constructs were verified by automated sequencing.

The wild-type and mutant GST-LBDs were expressed in BL21(DE3) cells. The total [^{3}H]E_2 binding activity in each extract was determined by saturation analysis using a controlled pore glass bead (CPG) assay (Greene et al., 1988). GST-LBD protein levels were also monitored by Western blotting with a monoclonal antibody to hERα (H222) to confirm that the mutant GST-LBDs bound E_2 with affinities comparable to the wild-type protein. Cleared extracts containing the GST-LBDs were incubated in buffer alone (50 mM Tris [pH 7.4], 150 mM NaCl, 2 mM EDTA, 1 mM DTT, 0.5% NP-40, and a protease inhibitor cocktail) or with 1 μM of either DES or OHT for 1 hr at 4°C. Extract aliquots containing 30 pmol of binding activity, based on the CPG assay, were then incubated with 10 μl glutathione-Sepharose-4B beads (Pharmacia) for 1 hr at 4°C. Beads were washed five times with 20 mM HEPES (pH 7.4), 400 mM NaCl, and 0.05% NP-40. ^{35}S-labeled GRIP1 was synthesized using the TNT Coupled Reticulocyte Lysate System (Promega) and pSG5-GRIP1 (a gift of M. Stallcup) as the template. Immobilized GST-LBDs were incubated for 2.5 hr with 2.5 μl aliquots of crude translation reaction mixture diluted in 300 μl of Tris-buffered saline (TBS). After five washes in TBS containing 0.05% NP-40, proteins were eluted by boiling the beads for 10 min in sample buffer. Bound ^{35}S-GRIP1 was visualized by fluorography following SDS-PAGE.

Crystallization and Data Collection

Crystals of the DES-hERα LBD-GRIP1 NR box II peptide complex were obtained by hanging drop vapor diffusion at 19°C–21°C. Prior to crystallization, the DES-LBD complex was incubated with a 2- to 4-fold molar excess of the GRIP1 NR box II peptide for 7–16 hr. Samples (2 μl) of this solution (4.3 mg/ml protein) were mixed with 2 μl of the reservoir buffer consisting of 25%–27% (w/v) PEG 4000, 90 mM Tris (pH 8.75–9.0), and 180 mM Na acetate and suspended over wells of the reservoir buffer. These crystals lie in the space group $P2_1$ with cell parameters a = 54.09 Å, b = 82.22 Å, c = 58.04 Å, and β = 111.34. Two molecules of the DES-LBD complex and of the coactivator peptide form the asymmetric unit. A crystal was transferred to a cryosolvent solution containing 25% (w/v) PEG 4000, 10% (w/v) ethylene glycol, 100 mM Tris (pH 8.5), 200 mM Na acetate, and 10 μM peptide and frozen in an N_2 stream at −170°C in a rayon loop. Diffraction data were measured at −170°C using the 300 mm MAR image plate at the Stanford Synchrotron Radiation Laboratory (SSRL) beamline 7-1 (λ = 1.08 Å).

Crystals of the OHT-hERα LBD complex were obtained by hanging drop vapor diffusion at 21°C–23°C. Samples (2 μl) of a solution containing 3.9 mg/ml complex and 2 μl of the reservoir solution containing 9% (w/v) PEG 8000, 6% (w/v) ethylene glycol, 50 mM HEPES (pH 6.7), and 200 mM NaCl were mixed and suspended over wells of the reservoir solution. These crystals lie in the space group $P6_122$ with cell parameters a = b = 58.24 Å and c = 277.47 Å. The asymmetric unit consists of a single LBD monomer; the dimer axis lies along a crystallographic 2-fold. A crystal was briefly incubated in a cryoprotectant solution consisting of 10% (w/v) PEG 8000, 25% (w/v) ethylene glycol, 50 mM HEPES (pH 7.0), and 200 mM NaCl and then flash frozen in liquid N_2 suspended in a rayon loop. Diffraction data were measured at −170°C at SSRL beamline 9-1 (λ = 0.98 Å) using a 345 MAR image plate.

The images of both data sets were processed with DENZO (Otwinowski and Minor, 1997), and both data sets were scaled with SCALEPACK (Otwinowski and Minor, 1997) using the default -3σ cutoff.

Structure Determination and Refinement

Our initial efforts to determine the structure of the DES-LBD-peptide complex utilized a low resolution (3.1 Å) data set (data not shown). The two LBDs in the asymmetric unit were located by molecular replacement in AMoRe (CCP4, 1994) using a partial polyalanine model of the human retinoic acid receptor γ LBD (Renaud et al., 1995) as the search probe (R= 58.2%, CC= 35.6% after placement of both monomers). Iterative cycles of 2-fold NCS averaging in DM (CCP4, 1994) interspersed with model building in MOLOC (Muller et al., 1988) and model refinement in REFMAC (Murshudov et al., 1997) (using tight NCS restraints and phases from 2-fold averaging) were used to quickly build a model of the LBD alone. For this procedure, MAMA (Kleywegt and Jones, 1994) was used for all mask manipulations, and PHASES (Furey and Swaminathan, 1990) and the CCP4 suite (CCP4, 1994) were used for the generation of structure factors and the calculation of weights.

At this point, although the DES-LBD complex model accounted for ~90% of the scattering matter in the asymmetric unit, refinement was being hampered by severe model bias. The OHT complex data set was then collected (Table 1). Starting with one of the monomers of the preliminary DES-LBD model as the search probe, molecular replacement in AMoRe was used to search for the location of LBD in this crystal form in both $P6_122$ and $P6_522$. A translation search in $P6_522$ yielded the correct solution (R= 53.8%, CC = 38.2%). In order to reduce model bias, DMMULTI (CCP4, 1994) was then used to project averaged density from the DES complex cell into the OHT complex cell. Using MOLOC, a model of the LBD was built into the resulting density. The model was refined initially in REFMAC and later with the simulated annealing, positional, and B-factor refinement protocols in X-PLOR (Brünger, 1996) using a maximum-likelihood target (Adams et al., 1997). Anisotropic scaling and a bulk solvent correction were used, and all B-factors were refined isotropically. Except for the R_{free} set (a random sampling consisting of 8% of the data set), all data between 41 and 1.9 Å (with no σ cutoff) were included. The final model consists of residues 306–551, the ligand, and 79 waters. According to PROCHECK (CCP4, 1994), 91.6% of all residues in the model are in the core regions of the Ramachandran plot and none are in the disallowed regions.

The high resolution data set of the DES-LBD NR box II peptide complex (Table 1) became available when the R_{free} of the OHT-LBD model was ~31%. Both monomers in the asymmetric unit of the DES complex crystal were relocated using AMoRe and the incompletely refined OHT-LBD model (with helix 12 and the loop between helices 11 and 12 removed) as the search model. The missing parts of the model were built, and the rest of the model was corrected using MOLOC and 2-fold averaged maps generated in DM. Initially, refinement was carried out with REFMAC, using tight NCS restraints. At later stages, the model was refined without NCS restraints using the simulated annealing, positional, and B-factor refinement protocols in X-PLOR and a maximum-likelihood target. All B-factors were refined isotropically, and anisotropic scaling and a bulk solvent correction were used. The R_{free} set contains a random sample of 6.5% of all data. In refinement, all data between 27 and 2.03 Å (with no σ cutoff) were used. The final model is composed of residues 305–549 of monomer A, residues 305–461 and 470–549 of monomer B, residues 687–697 of peptide A, residues 686–696 of peptide B, two ligand molecules, 147 waters, two carboxymethyl groups, and a chloride ion. According to PROCHECK, 93.7% of all residues in the model are in the core regions of the Ramachandran plot, and none are in the disallowed regions.

Illustrations

Figures 3C and 3D were created using GRASP (Nicholls et al., 1991). Figures 1, 2, 3A, 3B, 5, 6A, and 6D were generated using BOBSCRIPT (Esnouf, 1997) and rendered using Raster3D (Merritt and Anderson, 1994). Figure 4 was generated using LIGPLOT (Wallace et al., 1995), and Figures 6B and 6C were created using MidasPlus (Huang et al., 1991). Figure 1A depicts peptide B; all other illustrations of the coactivator peptide depict peptide A.

Acknowledgments

This work has been supported by funds from the U.S. Army Medical and Research Material Command grant DAMD 17-94-J-4228 (G. L. G.), the NCI Cancer Center Support grant P30 CA-14599 (G. L. G.), the Howard Hughes Medical Institute (D. A. A.), the NIGMS grant GM31627 (D. A. A.), the American Cancer Society grant BE61 (P. J. K.), and the NIH grant DK51083 (P. J. K.). A. K. S. was supported by a Howard Hughes Medical Institute Predoctoral Fellowship and a UCSF Chancellor's Fellowship. The Stanford Synchrotron Radiation Laboratory (SSRL) is funded by the Department of Energy, Office of Basic Energy Science. We thank H. Deacon, P. Foster, T. Mau, N. Sauter, and the staff of SSRL for assistance with data collection; C. Anderson, C. Hospelhorn, and T. McSherry for assistance with the biochemical experiments; and M. Butte, B. Darimont, R. Keenan, T. Mau, R. Wagner, and K. Yamamoto for comments on the manuscript. We would especially like to thank R. Wagner for contributions at the early stages of the project; D. Tanenbaum, Y. Wang, and P. Sigler for providing us their E_2-LBD coordinates in advance of publication; and H. Deacon and A. Derman for cheerful and extensive assistance with the figures and the manuscript, respectively.

Received September 1, 1998; revised November 16, 1998.

References

Adams, P.D., Pannu, N.S., Read, R.J., and Brünger, A.T. (1997). Cross-validated maximum likelihood enhances crystallographic simulated annealing refinement. Proc. Natl. Acad. Sci. USA *94*, 5018–5023.

Anzick, S.L., Kononen, J., Walker, R.L., Azorsa, D.O., Tanner, M.M., Guan, X.-Y., Sauter, G., Kallioniemi, O.-P., Trent, J.M., and Meltzer, P.S. (1997). AIB1, a steroid receptor coactivator amplified in breast and ovarian cancer. Science *277*, 965–968.

Beato, M., Herrlich, P., and Schutz, G. (1995). Steroid hormone receptors: many actors in search of a plot. Cell *83*, 851–857.

Berry, M., Metzger, D., and Chambon, P. (1990). Role of the two activating domains of the oestrogen receptor in the cell-type and

promoter-context dependent agonistic activity of the anti-oestrogen 4-hydroxytamoxifen. EMBO J. *9*, 2811–2818.

Bourguet, W., Ruff, M., Chambon, P., Gronemeyer, H., and Moras, D. (1995). Crystal structure of the ligand-binding domain of the human nuclear receptor RXR-α. Nature *375*, 377–382.

Brünger, A.T. (1996). X-PLOR Version 3.843 (New Haven, CT: Yale University).

Brzozowski, A., Pike, A., Dauter, Z., Hubbard, R., Bonn, T., Engstrom, O., Ohman, L., Greene, G., Gustafsson, J., and Carlquist, M. (1997). Molecular basis of agonism and antagonism in the oestrogen receptor. Nature *389*, 753–758.

CCP4. (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D *50*, 760–763.

Chen, H., Lin, R.J., Schiltz, R.L., Chakravarti, D., Nash, A., Nagy, L., Privalsky, M.L., Nakatani, Y., and Evans, R.M. (1997). Nuclear receptor coactivator ACTR is a novel histone acetyltransferase and forms a multimeric activation complex with P/CAF and CBP/p300. Cell *90*, 569–580.

Danielian, P., White, R., Lees, J., and Parker, M. (1992). Identification of a conserved region required for hormone dependent transcriptional activation by steroid hormone receptors. EMBO J. *11*, 1025–1033.

Ding, S., Anderson, C., Ma, H., Hong, H., Uht, R., Kushner, P., and Stallcup, M. (1998). Nuclear receptor-binding sites of coactivators glucocorticoid receptor interacting protein 1 (GRIP1) and steroid receptor coactivator 1 (SRC-1): multiple motifs with different binding specificities. Mol. Endocrinol. *12*, 302–313.

Esnouf, R.M. (1997). An extensively modified version of MolScript that includes greatly enhanced coloring capabilities. J. Mol. Graph. Model. *15*, 132–134.

Feng, W., Ribeiro, R.C., Wagner, R.L., Nguyen, H., Apriletti, J.W., Fletterick, R.J., Baxter, J.D., Kushner, P.J., and West, B.L. (1998). Hormone-dependent coactivator binding to a hydrophobic cleft on nuclear receptors. Science *280*, 1747–1749.

Furey, W., and Swaminathan, S. (1990). PA33. Am. Cryst. Assoc. Mtg. Abstr. *18*, 73.

Glass, C.K., Rose, D.W., and Rosenfeld, M.G. (1997). Nuclear receptor coactivators. Curr. Opin. Cell Biol. *9*, 222–232.

Gradishar, W.J., and Jordan, V.C. (1997). Clinical potential of new antiestrogens. J. Clin. Oncol. *15*, 840–852.

Grainger, D.J., and Metcalfe, J.C. (1996). Tamoxifen: teaching an old drug new tricks? Nat. Med. *2*, 381–385.

Greene, G., Nolan, C., Engler, J., and Jensen, E. (1980). Monoclonal antibodies to human estrogen receptor. Proc. Natl. Acad. Sci. USA *77*, 5115–5119.

Greene, G., Harris, K., Bova, R., Kinders, R., Moore, B., and Nolan, C. (1986). Purification of T47D human progesterone receptor and immunochemical characterization with monoclonal antibodies. Mol. Endocrinol. *2*, 714–726.

Grese, T.A., Sluka, J.P., Bryant, H.U., Cullinan, G.J., Glasebrook, A.L., Jones, C.D., Matsumoto, K., Palkowitz, A.D., Sato, M., Termine, J.D., et al. (1997). Molecular determinants of tissue selectivity in estrogen receptor modulators. Proc. Natl. Acad. Sci. USA *94*, 14105–14110.

Hanstein, B., Eckner, R., DiRenzo, J., Halachmi, S., Liu, H., Searcy, B., Kurokawa, R., and Brown, M. (1996). p300 is a component of an estrogen receptor coactivator complex. Proc. Natl. Acad. Sci. USA *93*, 11540–11545.

Heery, D., Kalkhoven, E., Hoare, S., and Parker, M. (1997). A signature motif in transcriptional co-activators mediates binding to nuclear receptors. Nature *387*, 733–736.

Hegy, G., Shackleton, C., Carlquist, M., Bonn, T., Engstrom, O., Sjoholm, P., and Witkowska, H. (1996). Carboxymethylation of the human estrogen receptor ligand-binding domain-estradiol complex: HPLC/ESMS peptide mapping shows that cysteine 447 does not react with iodoacetic acid. Steroids *61*, 367–373.

Henttu, P.M., Kalkhoven, E., and Parker, M.G. (1997). AF-2 activity and recruitment of steroid receptor coactivator 1 to the estrogen receptor depend on a lysine residue conserved in nuclear receptors. Mol. Cell. Biol. *17*, 1832–1839.

Hong, H., Kohli, K., Trivedi, A., Johnson, D.L., and Stallcup, M.R. (1996). GRIP1, a novel mouse protein that serves as a transcriptional coactivator in yeast for the hormone binding domains of steroid receptors. Proc. Natl. Acad. Sci. USA *93*, 4948–4952.

Horwitz, K.B., Jackson, T.A., Bain, D.L., Richer, J.K., Takimoto, G.S., and Tung, L. (1996). Nuclear receptor coactivators and corepressors. Mol. Endocrinol. *10*, 1167–1177.

Huang, C.C., Pettersen, E.F., Klein, T.E., Ferrin, T.E., and Langridge, R. (1991). Conic: a fast renderer for space-filling molecules with shadows. J. Mol. Graph. *9*, 230–236, 242.

Jordan, V.C. (1998). Antiestrogenic action of raloxifene and tamoxifen: today and tomorrow. J. Natl. Cancer Inst. *90*, 967–971.

Jordan, V.C., and Gosden, B. (1982). Importance of the alkylaminoethoxy side-chain for the estrogenic and antiestrogenic actions of tamoxifen and trioxifene in the immature rat uterus. Mol. Cell. Endocrinol. *27*, 291–306.

Kamei, Y., Xu, L., Heinzel, T., Torchia, J., Kurokawa, R., Gloss, B., Lin, S.C., Heyman, R.A., Rose, D.W., Glass, C.K., and Rosenfeld, M.G. (1996). A CBP integrator complex mediates transcriptional activation and AP-1 inhibition by nuclear receptors. Cell *85*, 403–414.

Kato, S., Endoh, H., Masuhiro, Y., Kitamoto, T., Uchiyama, S., Sasaki, H., Masushige, S., Gotoh, Y., Nishida, E., Kawashima, H., et al. (1995). Activation of the estrogen receptor through phosphorylation by mitogen-actived protein kinase. Science *270*, 1491–1494.

Kleywegt, G.J., and Jones, T.A. (1994). Halloween...masks and bones. In From First Map to Final Model, S. Bailey, R. Hubbard, and D. Waller, eds. (Warrington, England: SERC Daresbury Laboratory).

Korach, K. (1994). Insights from the study of animals lacking functional estrogen receptor. Science *266*, 1524–1527.

Kuiper, G.G., Carlsson, B., Grandien, K., Enmark, E., Haggblad, J., Nilsson, S., and Gustafsson, J.A. (1997). Comparison of the ligand binding specificity and transcript tissue distribution of estrogen receptors alpha and beta. Endocrinology *138*, 863–870.

Kumar, V., Green, S., Stack, G., Berry, M., Jin, J.-R., and Chambon, P. (1987). Functional domains of the human estrogen receptor. Cell *51*, 941–951.

Le Douarin, B., Nielsen, A.L., Garnier, J.M., Ichinose, H., Jeanmougin, F., Losson, R., and Chambon, P. (1996). A possible involvement of TIF1 alpha and TIF1 beta in the epigenetic control of transcription by nuclear receptors. EMBO J. *15*, 6701–6715.

Li, H., Gomes, P.J., and Chen, J.D. (1997). RAC3, a steroid/nuclear receptor-associated coactivator that is related to SRC-1 and TIF2. Proc. Natl. Acad. Sci. USA *94*, 8479–8484.

Merritt, E.A., and Anderson, W.F. (1994). Raster3D Version 2.0: a program for photorealistic molecular structures. Acta Crystallogr. D *50*, 219–220.

Moras, D., and Gronemeyer, H. (1998). The nuclear receptor ligand-binding domain: structure and function. Curr. Opin. Cell Biol. *10*, 384–391.

Muller, K., Amman, H.J., Doran, D.M., Gerber, P.R., Gubernator, K., and Schrepfer, G. (1988). MOLOC: A molecular modeling program. Bull. Soc. Chim. Belg. *97*, 655–667.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D *53*, 240–255.

Nicholls, A., Sharp, K., and Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins *11*, 281–296.

Norris, J.D., Fan, D., Stallcup, M.R., and McDonnell, D.P. (1998). Enhancement of estrogen receptor transcriptional activity by the coactivator GRIP-1 highlights the role of activation function 2 in determining estrogen receptor pharmacology. J. Biol. Chem. *273*, 6679–6688.

Onate, S.A., Tsai, S.Y., Tsai, M.-J., and O'Malley, B.W. (1995). Sequence and characterization of a coactivator for the steroid hormone receptor family. Science *270*, 1354–1357.

Otwinowski, Z., and Minor, W. (1997). Processing of x-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307–326.

Renaud, J., Rochel, N., Ruff, M., Vivat, V., Chambon, P., Gronemeyer,

H., and Moras, D. (1995). Crystal structure of the RAR-γ ligand-binding domain bound to all-trans retinoic acid. Nature *378*, 681–689.

Robertson, D.W., Katzenellenbogen, J.A., Hayes, J.R., and Katzenellenbogen, B.S. (1982). Antiestrogen basicity—activity relationships: a comparison of the estrogen receptor binding and antiuterotrophic potencies of several analogues of (Z)-1,2-diphenyl-1-[4-[2-(dimethylamino)ethoxy]phenyl]-1- butene (tamoxifen, Nolvadex) having altered basicity. J. Med. Chem. *25*, 167–171.

Seielstad, D., Carlson, K., Katzenellenbogen, J., Kushner, P., and Greene, G. (1995). Molecular characterization by mass spectrometry of the human estrogen receptor ligand-binding domain expressed in *Escherichia coli*. Mol. Endocrinol. *9*, 647–658.

Smigel, K. (1998). Breast cancer prevention trial shows major benefit, some risk. J. Natl. Cancer Inst. *90*, 647–648.

Smith, E., Boyd, J., Frank, G., Takahashi, H., Cohen, R., Specker, B., Williams, T., Lubahn, D., and Korach, K. (1994). Estrogen resistance caused by a mutation in the estrogen-receptor gene in a man. N. Engl. J. Med. *331*, 1056–1061.

Tanenbaum, D.M., Wang, Y., Williams, S.P., and Sigler, P.B. (1998). Crystallographic comparison of the estrogen and progesterone receptor's ligand binding domains. Proc. Natl. Acad. Sci. USA *95*, 5998–6003.

Tora, L., Mullick, A., Metzger, D., Ponglikitmongkol, M., Park, I., and Chambon, P. (1989). The cloned human oestrogen receptor contains a mutation which alters its hormone binding properties. EMBO J. *8*, 1981–1986.

Torchia, J., Rose, D., Inostroza, J., Kamei, Y., Westin, S., Glass, C., and Rosenfeld, M. (1997). The transcriptional co-activator p/CIP binds CBP and mediates nuclear-receptor function. Nature *387*, 677–684.

Tsai, M.J., and O'Malley, B.W. (1994). Molecular mechanisms of action of steroid/thyroid receptor superfamily members. Annu. Rev. Biochem. *63*, 451–486.

Voegel, J.J., Heine, M.J.S., Zechel, C., Chambon, P., and Gronemeyer, H. (1996). TIF2, a 160 kDa transcriptional mediator for the ligand-dependent activation function AF-2 of nuclear receptors. EMBO J. *15*, 3667–3675.

Voegel, J.J., Heine, M.J., Tini, M., Vivat, V., Chambon, P., and Gronemeyer, H. (1998). The coactivator TIF2 contains three nuclear receptor-binding motifs and mediates transactivation through CBP binding-dependent and -independent pathways. EMBO J. *17*, 507–519.

Wallace, A.C., Laskowski, R.A., and Thornton, J.M. (1995). LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. Protein Eng. *8*, 127–134.

Wrenn, C., and Katzenellenbogen, B. (1993). Structure-function analysis of the hormone binding domain of the human estrogen receptor by region-specific mutagenesis and phenotypic screening in yeast. J. Biol. Chem. *268*, 24089–24098.

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

Brookhaven Protein Data Bank Accession Numbers

Coordinates have been deposited with the PDB for the DES-ERα LBD-GRIP1 NR box II peptide complex (2ERD) and for the OHT-ERα LBD complex (2ERT).

Note Added in Proof

During the course of this work, we received a draft of a manuscript (Darimont, B.D., et al. [1998]. Structure and specificity of nuclear receptor-coactivator interactions. Genes Dev. *12*, 3343–3356) describing structural studies of the complex between TRβ and the GRIP1 NR box II peptide and biochemical studies of GRIP1 binding to TRβ and GR. We have included references to this work in the text as Darimont et al. (1998). The general features of the TRβ/GRIP1 NR box II peptide complex and the recently described PPARγ/SRC-1 peptide complex (Nolte, R.T., et al. [1998]. Ligand binding and co-activator assembly of the peroxisome proliferator-activated receptor-γ. Nature *395*, 137–143) are very similar to those of the ERα/GRIP1 NR box II peptide complex discussed here, suggesting that the mechanisms of NR box recognition are conserved across NRs.

Chapter 2

Structural Studies of an Estrogen Receptor α Agonist/Estrogen Receptor β Antagonist

Andrew K. Shiau*, Danielle Barstad†, John A. Katzenellenbogen‡, Benita S. Katzenellenbogen§, David A. Agard*, and Geoffrey L. Greene†

* The Howard Hughes Medical Institute
and the Department of Biochemistry and Biophysics
University of California at San Francisco
San Francisco, California 94143

† The Ben May Institute for Cancer Research
and the Department of Biochemistry and Molecular Biology
University of Chicago
Chicago, Illinois 60637

‡ Departments of Chemistry, §Molecular and Integrative Physiology, and §Cell and Structural Biology
University of Illinois at Urbana
Urbana, Illinois 61801

Introduction

Estrogens play a central role in mammalian physiology by regulating the differentiation, growth, and maintenance of a wide variety of tissues. For nearly thirty years, it was generally assumed that estrogens exerted their myriad effects by binding to and activating a single estrogen receptor (ER), the estrogen receptor α (ERα) (Green et al., 1986; Greene et al., 1986). This assumption was proven incorrect with the cloning and identification of a second estrogen receptor subtype, estrogen receptor β (ERβ) (Kuiper et al., 1996; Mosselman et al., 1996; Tremblay et al., 1997).

As might be expected, ERα and ERβ share many common structural and functional features. Both are members of the nuclear receptor (NR) superfamily of ligand-regulated transcription factors and exhibit the modular functional organization characteristic of most NRs: a central zinc finger DNA-binding domain (DBD) flanked by an amino-terminal domain and a carboxy-terminal ligand-binding domain (LBD). Although the sequences of the amino-terminal domains of ERα and ERβ share relatively little homology (~20% identity), the sequences of their DBDs and LBDs share 95% and 56% identity respectively (Green et al., 1986; Greene et al., 1986; Mosselman et al., 1996). The sequence similarities in these two domains allow both ERs to bind the "classical" inverted hexanucleotide repeat estrogen response element (ERE) (Pace et al., 1997) and to bind many estrogens and antiestrogens with comparable affinities (Kuiper et al., 1997).

Each of the ER isoforms also possesses at least two functionally separable transcriptional activation functions (AFs), AF-1 within the amino-terminal domain and AF-2 within the LBD. The activity of the AF-1 of each ER is regulated by growth factor-inducible phosphorylation by MAP kinases (Kato et al., 1995; Tremblay et al., 1999). In contrast, the activity of the AF-2 of each ER is stimulated by the binding of pure agonists, such as the endogenous estrogen 17β-estradiol (E2) and the synthetic stilbene estrogen diethylstilbestrol (DES), and blocked by the binding of pure antagonists, such as EM-800 and ICI-164,384 (Barkhem et al., 1998; Kumar et al., 1987; McInerney et al., 1998; Tremblay et al., 1998).

Despite these similarities, recent studies suggest that ERα and ERβ may play distinct roles in regulating gene expression *in vivo*. The two ERs have overlapping but distinct tissue distribution patterns; in rat, ERα is expressed in the uterus, testis, ovary, pituitary, kidney, epididymis, and adrenal glands and ERβ is expressed in the uterus, testis, ovary, prostate, lung, bladder, and brain (Kuiper et al., 1997). Consistent with these differences in localization, deletion of the ERα and ERβ genes in mice leads to substantially different phenotypes (Couse and Korach, 1999; Ogawa et al., 1999; Windahl et al., 1999).

The two ERs may also play distinct roles in the mechanism of action of certain therapeutics. Selective estrogen receptor modulators (SERMs) such as tamoxifen and raloxifene (RAL) are currently being used to treat a wide variety of diseases including osteoporosis, cardiovascular disease, and breast cancer (Gradishar and Jordan, 1997; Jordan, 1998). These compounds can function as either partial agonists or antagonists depending on the tissue and promoter

context. Tamoxifen, for example, acts as an antagonist in mammary tissue but as a partial agonist in endometrial, skeletal, and cardiovascular tissues (Grese et al., 1997). In mammalian cells, ERα and ERβ appear to respond differently to SERMs. RAL and 4-hydroxytamoxifen (OHT) trigger different responses from ERα and ERβ at ER-dependent AP-1 enhancer elements (Paech et al., 1997). In addition, at classical EREs, OHT and RAL act as partial agonists on ERα but as pure antagonists on ERβ (Barkhem et al., 1998; McInerney et al., 1998). These functional differences coupled with the distinct expression patterns of the two ERs may result in the complex, tissue-specific effects of SERMs. Compounds which act by novel mechanisms on the two ERs might therefore be potentially novel SERMs. This hypothesis has engendered a great deal of interest in subtype-selective ER ligands.

In a recent search for ER subtype-selective ligands, the R,R enantiomer of 5,11-*cis*-diethyl-5,6,11,12-tetrahydrochrysene-2,8-diol (THC) was identified as a compound with a modest preference for ERβ (K_d = 1.3 nM and 0.2 nM for full-length recombinant human ERα and human ERβ respectively) (Meyers et al., 1999; Sun et al., 1999). In transfection experiments in human endometrial cancer (HEC-1) cells and other cell lines, THC acts as a full agonist on ERα but has little, if any, effect on the transcriptional activity of ERβ (Meyers et al., 1999; Sun et al., 1999). However, THC is able to very effectively antagonize the effect of E2 on ERβ in a dose-dependent manner (IC_{50} < 10 nM) (Meyers et al., 1999; Sun et al., 1999).

Given its activity profile, THC may act on ERα and ERβ by mechanisms distinct from those of OHT and RAL. The overall transcriptional activity of each

ER results from the activities of both AF-1 and AF-2. For both ERα and ERβ, OHT and RAL act as AF-2 antagonists. The partial agonist character of these compounds on ERα in certain tissue and promoter contexts is thought to result from AF-1 activity (MacGregor and Jordan, 1998). OHT and RAL are thought to act as pure ERβ antagonists because the transcriptional activity of ERβ is largely dominated by AF-2 (Barkhem et al., 1998; McInerney et al., 1998). THC is similar to OHT and RAL in that it acts as a pure ERβ antagonist. In contrast, THC functions as a full agonist on ERα. This suggests that either THC is able to induce a hyperactive state of AF-1 (while acting as an AF-2 antagonist) or it is able to elicit activity from both AF-1 and AF-2.

Recent structural studies indicate that NR ligands modulate AF-2 activity by differentially affecting the structure of NR LBDs (Moras and Gronemeyer, 1998). The binding of an AF-2 agonist to an NR LBD stabilizes a conformation of the LBD which favors the binding of transcriptional coactivators (Darimont et al., 1998; Nolte et al., 1998; Chapter 1). These coactivators, which enhance ligand-dependent transcription of NRs, are thought to serve as the bridge between agonist-bound LBDs and the general transcriptional machinery (Glass et al., 1997). Members of the p160 family of coactivators, including SRC-1/NCoA1 (Kamei et al., 1996; Onate et al., 1995), GRIP1/TIF2/NCoA2 (Hong et al., 1996; Torchia et al., 1997; Voegel et al., 1996), and p/CIP/RAC3/ACTR/AIB1 (Anzick et al., 1997; Chen et al., 1997; Li et al., 1997; Torchia et al., 1997), recognize agonist-bound NR LBDs via short sequence motifs, LXXLL (where L is leucine and X is any amino acid) called NR boxes (Ding et al., 1998; Heery et al., 1997; Le Douarin et al., 1996; Torchia et al., 1997). These NR boxes form amphipathic α-

helices which recognize a hydrophobic groove on the surface of an agonist-bound LBD formed by residues from helices 3, 4, 5, and 12 (Darimont et al., 1998; Nolte et al., 1998; Chapter 1). In contrast, AF-2 antagonists such as OHT and RAL appear to stabilize an alternative conformation of the LBD that sterically precludes coactivator recognition (Brzozowski et al., 1997; Pike et al., 1999; Chapter 1). Helix 12 in the antagonist-bound structures of the ERα and ERβ LBDs occludes the coactivator recognition groove by mimicking the interactions of the NR box with the LBD (Brzozowski et al., 1997; Pike et al., 1999; Chapter 1). The recent structure of the ERβ LBD bound to the AF-2 partial agonist genistein (GEN) reveals that GEN binding drives helix 12 into yet a third conformation; helix 12 is bound over the ligand binding pocket in a position that only partially occludes the coactivator recognition surface (Pike et al., 1999). Depending on its actual mechanism of action, THC could induce either the agonist-bound, partial agonist-bound, or antagonist-bound conformations of the ERα LBD. However, given its effects on the transcriptional activity of ERβ, THC would be expected to induce the antagonist-bound conformation of the ERβ LBD.

To begin elucidating the mechanisms by which THC acts on the two ERs, we have used x-ray crystallography to determine the structure of the THC-ERα LBD-GRIP1 NR Box II peptide complex to 1.95 Å resolution and the structure of the THC-ERβ LBD complex to 2.85 Å resolution. The binding of THC to the ERα LBD stabilizes the agonist-bound conformation of helix 12. Remarkably, however, the binding of THC to the ERβ LBD stabilizes a conformation of helix 12 that most closely resembles that in the GEN-ERβ LBD complex (Pike et al.,

Results

Overall Structure of the THC-ERα LBD-GRIP1 NR Box II Peptide Complex

Although crystals of the THC-ERα LBD-GRIP1 NR Box II peptide complex and the DES-ERα LBD-GRIP1 NR Box II peptide complex were obtained under different conditions, the crystals of the two complexes lie in the same space group and have approximately the same cell dimensions (Experimental Procedures; Chapter 1). Not surprisingly, the structures of the two complexes are virtually identical (0.45 Å root mean square deviation on all Cα atoms). The asymmetric unit of the THC complex crystals contains two LBDs arranged in the same dimeric configuration that has been observed in all previous structures of the ERα LBD (Brzozowski et al., 1997; Tanenbaum et al., 1998; Chapter 1). Each LBD, a wedge-shaped molecule composed of a small beta hairpin and eleven to twelve helices arranged in three layers, is bound to a molecule of THC and a coactivator peptide (Figure 1A). Each THC molecule is completely enveloped within the lower subdomain of an LBD formed by residues from helices 3, 6, 7, 8, 11, and 12 and the S1/S2 hairpin (Figures 1A and 2A). Each coactivator peptide is bound in an α-helical conformation to a hydrophobic groove on the surface of an LBD formed by residues from helices 3, 4, 5, and 12 (Figure 1A).

THC Recognition by the ERα LBD

Overall, THC interacts with the ERα LBD in a manner analogous to the ways in which E2 and DES interact with the LBD (Brzozowski et al., 1997; Tanenbaum et al., 1998; Chapter 1). The A ring of THC is bound in approximately the same

position as the A rings of E2 and DES near helices 3 and 6. The A ring itself is recognized by the side chains of Ala 350, Leu 387, Leu 391, and Phe 404 and the A ring phenolic hydroxyl forms hydrogen bonds to the side chain carboxylate of Glu 353, to the guanidinium group of Arg 394, and to a buried water molecule (Figures 2A and 3). The A′ ring of THC interacts with the opposite end of the binding pocket near helices 7, 8, and 11 which interacts with the D ring of E2 and the A′ ring of DES. The A′ ring forms van der Waals contacts with Met 421, Ile 424, Gly 521, and Leu 525 and the phenolic hydroxyl forms a hydrogen bond with the side chain imidazole of His 524 (Figures 2A and 4).

THC, E2, and DES, however, differ in their ability to fill the binding pocket. In the E2-LBD complex, there are unoccupied cavities adjacent to the α face of the B ring and the β face of the C ring of E2 (Brzozowski et al., 1997; Tanenbaum et al., 1998). DES is able to fill these cavities with its two ethyl groups (Chapter 1). One of the ethyl groups of THC fills the cavity near the B ring by forming van der Waals contacts with Leu 391, Phe 404, Met 421, and Leu 428 (Figure 2A). The other ethyl group of THC points into the region of the binding pocket near the α face, instead of the β face, of the steroid C ring in the E2 complex. This ethyl group is, therefore, unable to fill the cavity near the C ring of E2 and instead forms van der Waals contacts with the side chains of both Leu 346 and Met 421 (Figure 2A). These packing interactions force the body of THC to be slightly angled (~10°) towards helix 12 relative to the body of E2.

Overall Structure of the THC-ERβ LBD Complex

The two ERβ LBD monomers in the THC-ERβ complex crystals do not interact to for the same dimer seen in the α complex. Instead, each of the two LBDs interact with symmetry-related molecules to form related but somewhat differently packed crystallographic trimers. The ERα LBD dimer has been shown to dissociate below pH 6.5, presumably through the protonation of one or more of the histidines at the dimer interface (Thole, 1994). Given that the residues involved in homodimerization are well conserved between the two isoforms (Green et al., 1986; Greene et al., 1986; Mosselman et al., 1996), the high salt and low pH (~4.5) conditions used to obtain crystals of the THC-ERβ LBD complex should also favor the dissociation of the ERβ LBD dimer into monomers. These monomers presumably interact to form trimers either prior to or during the process of during crystal growth.

Despite these differences in quaternary structure, each of the two ERβ LBD monomers in the asymmetric unit adopts the same overall fold as the ERα LBD. As has been described previously (Pike et al., 1999), the ERβ LBD is also a wedge-shaped molecule formed from three layers of eleven to twelve helices and a small beta hairpin (Figure 1B). The most notable feature of the conformation of the ERβ LBD bound to THC is that helix 12 is not positioned as it is in the ERα LBD-agonist complexes (Brzozowski et al., 1997; Chapter 1). Helix 12 is also not bound in the static region of the coactivator recognition groove as it is in the OHT- and RAL-ERα LBD complexes and the RAL-ERβ LBD complex (Brzozowski et al., 1997; Pike et al., 1999; Chapter 1). Instead, helix 12 in the THC-ERβ LBD complex (Figures 1B and 5) interacts with the rest of the LBD in a

manner most similar to that observed in the GEN-ERβ LBD complex (Pike et al., 1999).

The Conformation of Helix 12 in the THC-ERβ LBD Complex

Many of the same packing interactions dictate the positioning of helix 12 in both the THC (Figures 1B and 5) and the GEN-ERβ LBD complexes (Pike et al., 1999). As it does in the GEN complex, the side chain of Val 487, which forms van der Waals contacts with the side chains of Thr 299 and Leu 306, is inserted into the entrance of the ligand binding pocket formed by helices 3, 6, and 11 in the THC complex (Figure 5). The position of Val 487, which lies at the N-terminus of helix 12, forces the C-terminal end of helix 12 to project away from the body of the LBD (Figure 1B). As a result, the side chains of Leu 491, Met 494, and Leu 495, which mimic the interactions of the three NR box leucines with the static region of the coactivator recognition groove in the RAL-ERβ LBD complex (Pike et al., 1999), only partially occlude the coactivator binding site. The side chain of Leu 491, which lies almost directly above Val 487, packs against those of Leu 306, Glu 332, and Trp 335 (Figure 5). The side chain of Met 494 lies near helix 3 and packs against the side chains of Leu 306, Val 307, and Ile 310 (Figure 5). Leu 495 occupies approximately the same position as Leu 491 in the RAL-ERβ LBD complex (Pike et al., 1999) and packs against the side chains of Ile 310, Val 328, Leu 331, and Glu 332 (Figure 5). Overall, helix 12 in the THC and GEN-ERβ LBD complexes is tilted by ~25° relative to its orientation in the RAL-ERβ LBD complex (Pike et al., 1999).

Although it is positioned similarly, the conformation of helix 12 in the THC complex is not identical to that in the GEN complex. Helix 12 is slightly shorter in the THC complex (residues 487 to 497) than helix 12 in the GEN complex (residues 487 to 499). In addition, whereas the side chain of Tyr 488 is well resolved in electron density maps of the GEN complex (Pike et al., 1999), it is not visible in maps of the THC complex. As a result, there is no evidence in the THC complex for the hydrogen bond between the phenolic hydroxyl of Tyr 488 and the side chain carboxylate of Glu 332 that is seen in the GEN complex.

THC Recognition by the ERβ LBD

THC is recognized in a similar but not identical manner by the ERβ LBD as it is by the ERα LBD. Although the two ethyl groups of THC are bound in similar locations, they are not in the same conformation in the two complexes. Whereas the two ethyl groups of THC are positioned underneath the body of the tetrahydrochrysene in the α complex, they are rotated outwards in the β complex (Figures 3 and 4). One of these ethyl groups forms nonpolar contacts with Leu 343 and Phe 356 (equivalent to Leu 391 and Phe 404 in ERα) and fills the unoccupied cavity near the α face of the B ring of E2 (Figure 2B). In order to accommodate the extended conformation of this moiety, the body of the tetrahydrochrysene is shifted outwards towards the entrance of the binding pocket near helices 3, 6, and 11 relative to its position in the α complex (Figures 3 and 4).

As a result of this positional shift, THC, in the β complex, is only able to form a subset of the interactions observed in the α complex (Figures 2A and 2B).

As it is in the α complex, the phenolic A ring of THC, in the β complex, is recognized by the region of the binding pocket formed by residues from helices 3 and 6 and the beta hairpin. As described above, the A ring of THC forms van der Waals contacts with the side chains of four residues (Phe 404, Ala 350, Leu 387 and Leu 391) in the α complex. The A ring of THC, however, only interacts with the side chains of three of the equivalent residues (Phe 356, Ala 302, and Leu 339) in the β complex (Figures 2A, 2B, and 3). The side chain of Leu 343 in ERβ (equivalent to Leu 391 in ERα) forms van der Waals contacts with one of the ethyl groups of THC instead of interacting with the A ring (Figure 2B). In the α complex, the side chains of Glu 353 and Arg 394 form van der Waals contacts with the A ring of THC, which presumably enhance the alignment of the hydrogen bonds formed between these residues and the phenolic A ring hydroxyl (Figures 2A and 3). Because it is bound somewhat more distantly from the equivalent residues (Glu 305 and Arg 346) in the β complex, THC does not form nonpolar contacts with the side chains of these residues and the A ring hydroxyl only forms a single hydrogen bond with the side chain carboxylate of Glu 305 (Figure 2B and 3).

In the β complex, the A′ ring of THC is bound in the same region of the binding pocket as it is in the α complex and packs against Gly 472 and Leu 476 (equivalent to Gly 521 and Leu 525 in ERα) (Figures 2B and 4). However, the A′ ring phenolic hydroxyl is positioned 0.9 Å distant from its position in the α complex and is, therefore, unable to form a hydrogen bond with the imidazole of His 475 as it does with His 524 in the α complex (Figure 4).

Discussion

The Binding of GEN and THC and Helix 12 Positioning

In mammalian cells, GEN and THC exhibit different activities on ERβ; GEN acts as a partial agonist and THC acts as a pure antagonist. And despite the fact that no residues from helix 12 appear to interact directly with either GEN (Pike et al., 1999) or THC (Figure 2B), helix 12 is in the same conformation in the GEN and THC-ERβ LBD complexes. How can this structural data be reconciled with the biological activity of these compounds?

The simplest model that is consistent with all of the existing structural and functional data begins with the hypothesis that, in the unliganded or apo-ERβ LBD, helix 12, adopts the same conformation as that observed in the GEN and THC complexes. The GEN/THC-bound conformation of helix 12 should prevent interaction with coactivator (and hence result in transcriptional inactivity of the LBD) for two reasons. First, the residues from helix 12 that are predicted to form part of the coactivator recognition or AF-2 surface are inappropriately positioned. Second, helix 12 itself is bound such that it partially occludes the static region of the AF-2 surface formed by residues from helices 3, 4, and 5 (Chapter 1). If the position of helix 12 in the apo-LBD is in the GEN/THC-bound conformation, this would be consistent with the observation that AF-2 is largely transcriptionally silent in the absence of ligand (Barkhem et al., 1998; McInerney et al., 1998). But in the absence of ligand, ERβ actually displays low but significant AF-2 mediated transcriptional activity (Barkhem et al., 1998). Hence, helix 12 in the apo-LBD must be in equilibrium between the GEN/THC-bound

conformation and the agonist-bound conformation (as seen in the E2 and DES-ERα LBD complexes), which is properly aligned for coactivator recognition with the GEN/THC-bound conformation being heavily favored over the agonist-bound conformation (Figure 6).

The binding of ligands to the LBD would serve either to perturb or change the nature of this equilibrium. Agonist binding, by stabilizing the formation of interactions in and around the binding pocket (Chapter 1), would promote AF-2 activity by shifting the equlibrium to favor the conformation of helix 12 that allows coactivator binding (Figure 6). The large basic side chains of OHT and RAL exit the binding pocket through the entrance formed by residues from helices 3, 6, and 11 and sterically preclude either the GEN/THC-bound or agonist-bound conformations of helix 12 which both involve helix 12 covering the entrance to the binding pocket (Figure 5) (Pike et al., 1999). In addition, antagonists like OHT and RAL induce structural distortions in and around the ligand binding pocket which disfavor the formation of interactions observed in agonist-bound structures of the LBD. In particular, these distortions cause helix 11 to be significantly shorter in the antagonist complexes than in the agonist complexes (Chapter 1). Therefore, the binding of these antagonists heavily favor a third and inactive conformation of the helix 12 in which helix 12 is bound to the static region of the coactivator binding groove, mimicking the interactions formed by an NR box with the rest of the LBD (Figure 6) (Brzozowski et al., 1997; Pike et al., 1999; Chapter 1). This conformation of helix 12 makes the binding of coactivators to the OHT and RAL complexes even less favorable than binding to the apo-LBD and is consistent with the inverse agonism (or reduction in basal

transcriptional activity) exhibited by these molecules in transcriptional assays (Barkhem et al., 1998).

Partial agonists, like GEN, would be ligands which are able to stabilize many, but not all, of the requisite interactions in and around the binding pocket that favor the agonist-bound or active conformation of helix 12. Although GEN appears to form many of the interactions with the ERβ ligand binding pocket that E2 and DES do with the ERα ligand binding pocket, helix 11 is shorter the GEN-ERβ LBD complex than in the agonist-bound ERα LBD complexes (Pike et al., 1999). Helix 11 in the GEN complex ends at His 475 as it does in the RAL-ERβ LBD complex (the length of helix 11 in the agonist-bound complexes precludes the GEN/THC-bound conformation as it does the OHT/RAL-bound conformation) (Chapter 1). So partial agonists, like GEN, may only partially shift the equilibrium in favor of the active conformation (Figure 6). In summary, the binding of coactivators to the GEN-ERβ LBD would be predicted to be much more favorable than binding to the OHT and RAL-LBD complexes, somewhat more favorable than binding to the apo-LBD, and less favorable than binding to the E2 or DES-LBD complexes. This would be consistent with the fact that the transcriptional activity of the GEN-bound ERβ is intermediate between that of pure-agonist bound ERβ and the apo-ERβ.

This model of partial agonist behavior is entirely consistent with structural and biological data on the peroxisome proliferator-activated receptor γ (PPARγ), the only nuclear receptor for which the apo and liganded structures of the LBD have been determined (Nolte et al., 1998). In the apo-PPARγ LBD structure, helix 12 is loosely bound to the LBD in a conformation that is

inconsistent with coactivator recognition (Nolte et al., 1998). When bound to the PPARγ LBD, PPARγ agonists, such as rosiglitazone, form a series of indirect hydrogen bonds with the phenolic hydroxyl of Tyr 473 from helix 12 (Nolte et al., 1998). This hydrogen bond network stabilizes a conformation of helix 12 that allows the binding of coactivators. The recent crystal structure of PPARγ with the weak partial agonist, GW0072, reveals that this ligand is able to fill the binding pocket without forming this hydrogen bond network. The conformation of helix 12 in this complex most closely resembles that in the apo-PPARγ LBD structure (Oberfield et al., 1999). The GW0072-PPARγ LBD complex is, however, able to interact with coactivators such as CBP and SRC-1 in mammalian two-hybrid experiments but to a much lesser extent than the rosiglitazone-PPARγ LBD complex (Oberfield et al., 1999). This demonstrates that the binding of GW0072 does not preclude the active conformation of the PPARγ LBD, it simply does not favor this conformation as much as the binding of a full agonist. For PPARγ, then, partial agonists are simply molecules that can bind to the LBD but only partially shift the helix 12 conformational equilibrium to favor the agonist-bound conformation.

The limiting case of a partial agonist would be a ligand, which could form some interactions with the ligand binding pocket but could not stabilize enough of the appropriate interactions in and around the binding pocket to shift the conformational equilibrium of helix 12 towards the active conformation at all. As long as it were able to bind to the LBD with high affinity, this ligand could potentially antagonize the effects of agonists on the LBD. THC might well be such a ligand for the ERβ (Figure 6). THC acts like GEN in that it stabilizes a

conformation of the LBD in which helix 11 ends at His 475. However, THC forms even fewer interactions with the ERβ LBD than GEN. Whereas, GEN forms hydrogen bonds with the side chains of Glu 305, Arg 346, and His 475 (Pike et al., 1999), THC is only able to form a single hydrogen bond with the γ-carboxylate of Glu 305. THC also forms fewer van der Waals contacts with the binding pocket than GEN (Figure 2B) (Pike et al., 1999).

The diffraction characteristics of the crystals of the THC complex also support the idea that THC interacts less optimally with the ERβ LBD than GEN. The relatively limited resolution of the data measured from these crystals (2.85 Å) largely results from the fact that the average intensity of the diffraction data falls off very rapidly as a function of resolution (Wilson B factor ~100 $Å^2$). By contrast, the crystals of the GEN complex do not suffer from the same problem and diffract to significantly higher resolution (1.8 Å) (Pike et al., 1999). In addition, diffraction images from the crystals of the THC complex show extensive evidence of thermal diffuse scatter. Both of these properties are suggestive that the molecules in the crystal lattice are undergoing large-scale breathing motions (Chang et al., 1998; Clarage and Phillips Jr., 1997). This would be expected if the LBD were not well packed around THC. The LBD, however, does appear to be well packed around GEN (Pike et al., 1999).

If the many interactions formed by GEN with the binding pocket are only able to partially shift the helix 12 equilibrium towards the agonist-bound conformation, then the even fewer interactions made by THC may not be able to influence the equilibrium at all. As a result, with respect to coactivator binding, the THC complex would be predicted to have a significantly lower affinity for

coactivator than the GEN complex and perhaps one that was comparable to those of the OHT and RAL complexes. This would be consistent with the effects of THC on the transcriptional activity of ERβ.

THC Isoform Specific Interactions and "Antagonism Without Side Chains"

THC is clearly able to induce the agonist-bound conformation of helix 12 when it binds to the ERα LBD. A comparison of the ERα and ERβ complexes provides some clues as to the nature of the interactions that THC fails to form with the ERβ ligand binding pocket that result in its inability to shift the conformational equilibrium of helix 12 in ERβ. Of the approximately twenty-two residues which line the ligand binding pocket, there are only two residues which are not identical in the two ER isoforms: Leu 384 (ERα)/Met 336 (ERβ) and Met 421(ERα)/Ile 373(ERβ) (Figure 7). Given that neither Leu 384 nor Met 336 directly contact THC in their respective complexes, neither of these two residues would be expected to contribute substantially to the isoform specific interactions with THC (Figures 2A and 2B). By contrast, Met 421 in the ERα complex forms van der Waals contacts with the A′ ring of and the two ethyl groups of THC, whereas Ile 373 in the ERβ complex does not interact at all with the ligand (Figures 2A and 2B). By forming the appropriate packing interactions, Met 421 aids in aligning the A′ ring such that the phenolic hydroxyl forms a hydrogen bond with the imidazole side chain of His 524 (Figure 4). In addition, Met 421 and several other binding pocket residues, including Met 343, His 524, and Met 528, form a series of packing interactions between each other and the ligand. These cooperative packing interactions stabilize a length of helix 11 that favors

the agonist-bound conformation of helix 12 (Chapter 1). In ERβ, the sidechain of Ile 373 is apparently either too small or not flexible enough to adopt the proper conformation necessary to interact with the ligand and Met 295, His 475, and Met 479 (equivalent to Met 343, His 524, and Met 528 in ERα). Hence, THC binding does not stabilize secondary structure to the same extent in ERβ as it does in ERα. Helix 11 as well as the two strands of the β-hairpin and helix 3 are shorter in the ERβ complex than they are in the ERα complex (Figure 8). The importance of Met 421 and Ile 373 to the isoform-specific effects of THC could be readily determined by analyzing the ligand binding and transcriptional properties of a mutant ERα in which Met 421 was mutated to an isoleucine and a mutant ERβ in which Ile 373 was mutated to a methionine.

In conclusion, the structures of THC bound to the LBDs of the two ERs reveals the mechanisms by which this compound acts. THC is a ERα AF-2 agonist and an ERβ AF-2 antagonist. THC antagonizes ERβ in a manner very different from OHT and RAL. OHT and RAL antagonize the receptor through side chain effects and the introduction of structural distortions in and around the binding pocket. In contrast, THC acts as a very weak partial agonist by filling the ligand binding pocket of the ERβ LBD suboptimally. The utilization of subtype-specific interactions in ligand design should allow for creation of new compounds that act differently on the two ERs and possess novel therapeutic properties.

Experimental Procedures

Protein Expression and Purification

The human ERα LBD (residues 297-554) was expressed, carboxymethylated, and purified as previously described (Chapter 1). The human ERβ LBD (residues 256-505) was expressed as a N-terminally hexahistidine-tagged fusion protein in BL21(DE3)pLysS harboring a derivative of pET-15b. Bacterial lysates were applied to an estradiol-Sepharose column and the bound ERβ LBD was carboxymethylated with 20 mM iodoacetic acid. Protein was eluted with 3 x 10^{-5} M THC in ~50 ml of 20 mM Tris·Cl, 1M Urea, and 10% DMF pH 8.1. The ERβ LBD was further purified by ion exchange chromatography (Resource Q, Pharmacia). Protein samples were analyzed by SDS-PAGE, native PAGE, and electrospray ionization mass spectrometry.

Crystallization and Data Collection

Crystals of the THC-ERα LBD-GRIP1 NR box II peptide complex were prepared by hanging drop vapor diffusion at 19-21°C . Prior to crystallization, the THC-ERα LBD complex was incubated overnight with a four-fold molar excess of the GRIP1 NR box II peptide. Samples (0.5 μL) of this solution (5.0 mg/mL protein) were mixed with 3.5 μL of reservoir buffer consisting of 16% (w/v) PEG 4000, 53 mM Tris·Cl (pH 8.8) and 50 mM $MgCl_2$ and suspended over wells containing 800 μL of the reservoir buffer. These crystals lie in the spacegroup $P2_1$ with cell dimensions a=54.55 Å, b=82.60 Å, c=59.04 Å and β=111.53°. Two molecules of the THC-LBD and two of the coactivator peptide form the asymmetric unit. A

crystal was transferred to a cryosolvent solution containing 20% (w/v) PEG 4000, 15% (w/v) ethylene glycol, 100 mM Tris (pH 8.6), 100 mM $MgCl_2$ and frozen in an N_2 stream in a rayon loop. Diffraction data were measured at -170°C at beamline 5.0.2 at ALS using a ADSC Quantum 4 CCD camera at a wavelength of 1.10 Å. Images were processed with DENZO (Otwinowski and Minor, 1997) and the integrated intensities were scaled with SCALEPACK (Otwinowski and Minor, 1997) using the default -3σ cutoff.

Crystals of the THC-ERβ LBD complex were obtained by hanging drop vapor diffusion at 21-23°C. Samples (2 µL) of a solution of the complex (4.8 mg/mL) were combined with 2 µL samples of a reservoir solution containing 1.5-1.75 M $(NH_4)_2SO_4$ and 100 mM NaOAc pH 4.8-5.2 and suspended over wells containing 800 µL of reservoir solution. Crystals formed within 1-2 days. These crystals belong to the space group R3 with cell parameters a=b=99.14 Å and c=193.38 Å (hexagonal setting). The asymmetric unit contains two LBD monomers. Prior to data collection, a single crystal was transferred to a stabilizing solution (1.8 M $(NH_4)_2SO_4$, 100 mM NaCl, 100 mM NaOAc pH 4.5, and 10 µM THC). The crystal was then sequentially transferred at 30 minute intervals through a series of solutions consisting of the stabilizing solution supplemented with increasing concentrations of ethylene glycol (1% increments) to a final concentration of 15%. The crystal was then flash frozen in an N_2 stream in a rayon loop. Diffraction data were measured at -170°C at beamline 5.0.2 at ALS using a ADSC Quantum 4 CCD camera at a wavelength of 1.07 Å. Images were processed with DENZO and the integrated intensities were scaled with SCALEPACK using the default -3σ cutoff.

Structure Determination and Refinement

The side chains of Met 343, Leu 346, Glu 353, Leu 384, Leu 387, Met 388, Leu 391, Arg 394, Phe 404, Met 421, Ile 424, Leu 428, His 524, Leu 525, Met 528, and Leu 540 of both monomers were converted to alanine and the ligands, ions, carboxymethyl groups and waters were removed from the DES complex model (Chapter 1). The model was subject to rigid body refinement in REFMAC (Murshudov et al., 1997). The missing parts of the model were built and the rest of the model was corrected using MOLOC (Muller et al., 1988) and two-fold averaged maps generated in DM (CCP4, 1994). All masks were generated using MAMA (Kleywegt and Jones, 1994). Initially, positional refinement was carried out with REFMAC. At later stages, the model was refined using the simulated annealing, positional and B-factor refinement protocols in CNS (Adams et al., 1997) and a maximum-likelihood target. All B-factors were refined isotropically and anisotropic scaling and a bulk solvent correction were used. The R_{free} set contains a random sample of 6.5% of all data. In refinement, all data between 47 and 1.95 Å (with no σ cutoff) were used. The current model is composed of residues 305-460 and 469-547 of monomer A, residues 305-461 and 472-547 of monomer B, residues 687-697 of peptide A, residues 686-696 of peptide B, two ligand molecules, and two water molecules.

The intensities within the data set fall off very rapidly as a function of resolution (Wilson B factor ~100 $Å^2$ calculated from 4 to 2.85 Å). In order to increase the contribution of the higher resolution terms the data was sharpened with a correction factor of –50 $Å^2$. All subsequent manipulations were performed using this sharpened data.

The two LBDs in the asymmetric unit were located by molecular replacement in AMoRe (CCP4, 1994) and TFFC (CCP4, 1994). The search model was constructed by overlapping the models of five ERα LBD complexes (PDB Accession Numbers 1A52, 1ERE, 1ERR, 3ERD, and 3ERT) and setting the occupancies of each model to 0.2 (R= 55.3%, CC= 55.9% after placement of both monomers). The model was then rebuilt using MOLOC and two-fold averaged maps generated using DM (MAMA was used for all mask manipulations). Initially, refinement was carried out with using tight NCS restraints with REFMAC and later with the simulated annealing, positional and B-factor refinement protocols in CNS and a maximum-likelihood target. All B-factors were refined isotropically and anisotropic scaling and a bulk solvent correction were used. The R_{free} set contains a random sample of 5% of all data. In refinement, all data between 49 and 2.85 Å (with no σ cutoff) were used. The current model is composed of residues 259 to 501 of both monomers A and B, two ligand molecules, and two water molecules.

Illustrations

Figures 1, 3, 4, 5, and 8 were generated using BOBSCRIPT (Esnouf, 1997) and rendered using Raster3D (Merritt and Anderson, 1994). Figure 2 was generated using LIGPLOT (Wallace et al., 1995).

References

Adams, P. D., Pannu, N. S., Read, R. J., and Brunger, A. T. (1997). Cross-validated maximum likelihood enhances crystallographic simulated annealing refinement. Proc. Natl. Acad. Sci. U S A *94*, 5018-23.

Anzick, S. L., Kononen, J., Walker, R. L., Azorsa, D. O., Tanner, M. M., Guan, X.-Y., Sauter, G., Kallioniemi, O.-P., Trent, J. M., and Meltzer, P. S. (1997). AIB1, a steroid receptor coactivator amplified in breast and ovarian cancer. Science *277*, 965-968.

Barkhem, T., Carlsson, B., Nilsson, Y., Enmark, E., Gustafsson, J., and Nilsson, S. (1998). Differential response of estrogen receptor alpha and estrogen receptor beta to partial estrogen agonists/antagonists. Mol. Pharmacol. *54*, 105-12.

Brzozowski, A., Pike, A., Dauter, Z., Hubbard, R., Bonn, T., Engstrom, O., Ohman, L., Greene, G., Gustafsson, J., and Carlquist, M. (1997). Molecular basis of agonism and antagonism in the oestrogen receptor. Nature *389*, 753-758.

CCP4 (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D *50*, 760-763.

Chang, G., Spencer, R. H., Lee, A. T., Barclay, M. T., and Rees, D. C. (1998). Structure of the MscL homolog from Mycobacterium tuberculosis: a gated mechanosensitive ion channel. Science *282*, 2220-6.

Chen, H., Lin, R. J., Schiltz, R. L., Chakravarti, D., Nash, A., Nagy, L., Privalsky, M. L., Nakatani, Y., and Evans, R. M. (1997). Nuclear receptor coactivator ACTR is a novel histone acetyltransferase and forms a multimeric activation complex with P/CAF and CBP/p300. Cell *90*, 569-80.

Clarage, J. B., and Phillips Jr., G. N. (1997). Analysis of Diffuse Scattering and Relation to Molecular Motion. Methods Enzymol. *277*, 407-432.

Couse, J. F., and Korach, K. S. (1999). Estrogen receptor null mice: what have we learned and where will they lead us? Endocr. Rev. *20*, 358-417.

Darimont, B. D., Wagner, R. L., Apriletti, J. W., Stallcup, M. R., Kushner, P. J., Baxter, J. D., Fletterick, R. J., and Yamamoto, K. R. (1998). Structure and specificity of nuclear receptor-coactivator interactions. Genes Dev. *12*, 3343-56.

Ding, S., Anderson, C., Ma, H., Hong, H., Uht, R., Kushner, P., and Stallcup, M. (1998). Nuclear receptor-binding sites of coactivators glucocorticoid receptor interacting protein 1 (GRIP1) and steroid receptor coactivator 1 (SRC-1): Multiple motifs with different binding specificities. Mol. Endocrinol. *12*, 302-313.

Esnouf, R. M. (1997). An extensively modified version of MolScript that includes greatly enhanced coloring capabilities. J. Mol. Graph. Model. *15*, 132-4, 112-3.

Glass, C. K., Rose, D. W., and Rosenfeld, M. G. (1997). Nuclear receptor coactivators. Curr. Opin. Cell Biol. *9*, 222-32.

Gradishar, W. J., and Jordan, V. C. (1997). Clinical potential of new antiestrogens. J. Clin. Oncol. *15*, 840-52.

Green, S., Walter, P., Kumar, V., Krust, A., Bornert, J. M., Argos, P., and Chambon, P. (1986). Human oestrogen receptor cDNA: sequence, expression and homology to v-erb-A. Nature *320*, 134-9.

Greene, G. L., Gilna, P., Waterfield, M., Baker, A., Hort, Y., and Shine, J. (1986). Sequence and expression of human estrogen receptor complementary DNA. Science *231*, 1150-4.

Grese, T. A., Sluka, J. P., Bryant, H. U., Cullinan, G. J., Glasebrook, A. L., Jones, C. D., Matsumoto, K., Palkowitz, A. D., Sato, M., Termine, J. D., Winter, M. A., Yang, N. N., and Dodge, J. A. (1997). Molecular determinants of tissue selectivity in estrogen receptor modulators. Proc. Natl. Acad. Sci. U S A *94*, 14105-10.

Heery, D., Kalkhoven, E., Hoare, S., and Parker, M. (1997). A signature motif in transcriptional co-activators mediates binding to nuclear receptors. Nature *387*, 733-736.

Hong, H., Kohli, K., Trivedi, A., Johnson, D. L., and Stallcup, M. R. (1996). GRIP1, a novel mouse protein that serves as a transcriptional coactivator in yeast for the hormone binding domains of steroid receptors. Proc. Natl. Acad. Sci. USA *93*, 4948-4952.

Jordan, V. C. (1998). Antiestrogenic action of raloxifene and tamoxifen: today and tomorrow. J. Natl. Cancer Inst. *90*, 967-71.

Kamei, Y., Xu, L., Heinzel, T., Torchia, J., Kurokawa, R., Gloss, B., Lin, S. C., Heyman, R. A., Rose, D. W., Glass, C. K., and Rosenfeld, M. G. (1996). A CBP integrator complex mediates transcriptional activation and AP-1 inhibition by nuclear receptors. Cell *85*, 403-14.

Kato, S., Endoh, H., Masuhiro, Y., Kitamoto, T., Uchiyama, S., Sasaki, H., Masushige, S., Gotoh, Y., Nishida, E., Kawashima, H., Metzger, D., and Chambon, P. (1995). Activation of the estrogen receptor through phosphorylation by mitogen-actived protein kinase. Science *270*, 1491-1494.

Kleywegt, G. J., and Jones, T. A. (1994). Halloween...masks and bones. In From First Map to Final Model, S. Bailey, R. Hubbard and D. Waller, eds. (Warrington, England: SERC Daresbury Laboratory).

Kuiper, G., Enmark, E., Pelto-Huikko, M., Nilsson, S., and Gustafsson, J. (1996). Cloning of a novel estrogen receptor expressed in rat prostate and ovary. Proc. Natl. Acad. Sci. *93*, 5925-5930.

Kuiper, G. G., Carlsson, B., Grandien, K., Enmark, E., Haggblad, J., Nilsson, S., and Gustafsson, J. A. (1997). Comparison of the ligand binding specificity and transcript tissue distribution of estrogen receptors alpha and beta. Endocrinology *138*, 863-70.

Kumar, V., Green, S., Stack, G., Berry, M., Jin, J.-R., and Chambon, P. (1987). Functional domains of the human estrogen receptor. Cell *51*, 941-951.

Le Douarin, B., Nielsen, A. L., Garnier, J. M., Ichinose, H., Jeanmougin, F., Losson, R., and Chambon, P. (1996). A possible involvement of TIF1 alpha and TIF1 beta in the epigenetic control of transcription by nuclear receptors. EMBO J. *15*, 6701-15.

Li, H., Gomes, P. J., and Chen, J. D. (1997). RAC3, a steroid/nuclear receptor-associated coactivator that is related to SRC-1 and TIF2. Proc. Natl. Acad. Sci. U S A *94*, 8479-84.

MacGregor, J. I., and Jordan, V. C. (1998). Basic guide to the mechanisms of antiestrogen action. Pharmacol. Rev. *50*, 151-96.

McInerney, E. M., Weis, K. E., Sun, J., Mosselman, S., and Katzenellenbogen, B. S. (1998). Transcription activation by the human estrogen receptor subtype beta (ER beta) studied with ER beta and ER alpha receptor chimeras. Endocrinology *139*, 4513-22.

Merritt, E. A., and Anderson, W. F. (1994). Raster3D Version 2.0: A program for photorealistic molecular structures. Acta Crystallogr. D *50*, 219-220.

Meyers, M. J., Sun, J., Carlson, K. E., Katzenellenbogen, B. S., and Katzenellenbogen, J. A. (1999). Estrogen receptor subtype-selective ligands: asymmetric synthesis and biological evaluation of cis- and trans-5,11-dialkyl-5,6,11, 12- tetrahydrochrysenes. J. Med. Chem. *42*, 2456-68.

Moras, D., and Gronemeyer, H. (1998). The nuclear receptor ligand-binding domain: structure and function. Curr. Opin. Cell Biol. *10*, 384-91.

Mosselman, S., Polman, J., and Dijkema, R. (1996). ER beta: identification and characterization of a novel human estrogen receptor. FEBS Lett. *392*, 49-53.

Muller, K., Amman, H. J., Doran, D. M., Gerber, P. R., Gubernator, K., and Schrepfer, G. (1988). MOLOC: A molecular modeling program. Bull. Soc. Chim. Belg. *97*, 655-667.

Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D *53*, 240-255.

Nolte, R. T., Wisely, G. B., Westin, S., Cobb, J. E., Lambert, M. H., Kurokawa, R., Rosenfeld, M. G., Willson, T. M., Glass, C. K., and Milburn, M. V. (1998). Ligand binding and co-activator assembly of the peroxisome proliferator-activated receptorγ. Nature *395*, 137-43.

Oberfield, J. L., Collins, J. L., Holmes, C. P., Goreham, D. M., Cooper, J. P., Cobb, J. E., Lenhard, J. M., Hull-Ryde, E. A., Mohr, C. P., Blanchard, S. G., Parks, D. J., Moore, L. B., Lehmann, J. M., Plunket, K., Miller, A. B., Milburn, M. V., Kliewer, S. A., and Willson, T. M. (1999). A peroxisome proliferator-activated receptor gamma ligand inhibits adipocyte differentiation. Proc. Natl. Acad. Sci. U S A *96*, 6102-6.

Ogawa, S., Chan, J., Chester, A. E., Gustafsson, J. A., Korach, K. S., and Pfaff, D. W. (1999). Survival of reproductive behaviors in estrogen receptor beta gene-deficient (betaERKO) male and female mice. Proc. Natl. Acad. Sci. U S A *96*, 12887-92.

Onate, S. A., Tsai, S. Y., Tsai, M.-J., and O'Malley, B. W. (1995). Sequence and characterization of a coactivator for the steroid hormone receptor family. Science *270*, 1354-1357.

Otwinowski, Z., and Minor, W. (1997). Processing of x-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307-326.

Pace, P., Taylor, J., Suntharalingam, S., Coombes, R. C., and Ali, S. (1997). Human estrogen receptor beta binds DNA in a manner similar to and dimerizes with estrogen receptor alpha. J. Biol. Chem. *272*, 25832-8.

Paech, K., Webb, P., Kuiper, G., Nilsson, S., Gustafsson, J., Kushner, P., and Scanlan, T. (1997). Differential Ligand Activation of Estrogen Receptors ERα and ERβ at AP1 Sites. Science *277*, 1508-1510.

Pike, A. C., Brzozowski, A. M., Hubbard, R. E., Bonn, T., Thorsell, A. G., Engstrom, O., Ljunggren, J., Gustafsson, J. A., and Carlquist, M. (1999). Structure of the ligand-binding domain of oestrogen receptor beta in the presence of a partial agonist and a full antagonist. EMBO J. *18*, 4608-18.

Sun, J., Meyers, M. J., Fink, B. E., Rajendran, R., Katzenellenbogen, J. A., and Katzenellenbogen, B. S. (1999). Novel ligands that function as selective estrogens or antiestrogens for estrogen receptor-alpha or estrogen receptor-beta. Endocrinology *140*, 800-4.

Tanenbaum, D. M., Wang, Y., Williams, S. P., and Sigler, P. B. (1998). Crystallographic comparison of the estrogen and progesterone receptor's ligand binding domains. Proc. Natl. Acad. Sci. U S A *95*, 5998-6003.

Thole, H. H. (1994). The side chains responsible for the dimerization of the estradiol receptor by ionic bonds are lost in a 17 kDa fragment extending downstream from aa 303. J Steroid Biochem. Mol. Biol. *48*, 463-6.

Torchia, J., Rose, D., Inostroza, J., Kamel, Y., Westin, S., Glass, C., and Rosenfeld, M. (1997). The transcriptional co-activator p/CIP binds CBP and mediates nuclear-receptor function. Nature *387*, 677-684.

Tremblay, A., Tremblay, G. B., Labrie, C., Labrie, F., and Giguere, V. (1998). EM-800, a novel antiestrogen, acts as a pure antagonist of the transcriptional functions of estrogen receptors alpha and beta. Endocrinology *139*, 111-8.

Tremblay, A., Tremblay, G. B., Labrie, F., and Giguere, V. (1999). Ligand-independent recruitment of SRC-1 to estrogen receptor beta through phosphorylation of activation function AF-1. Mol. Cell *3*, 513-9.

Tremblay, G. B., Tremblay, A., Copeland, N. G., Gilbert, D. J., Jenkins, N. A., Labrie, F., and Giguere, V. (1997). Cloning, chromosomal localization, and functional analysis of the murine estrogen receptor beta. Mol. Endocrinol. *11*, 353-65.

Voegel, J. J., Heine, M. J. S., Zechel, C., Chambon, P., and Gronemeyer, H. (1996). TIF2, a 160 kDa transcriptional mediator for the ligand-dependent activation function AF-2 of nuclear receptors. EMBO J. *15*, 3667-3675.

Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1995). LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. Protein Eng. *8*, 127-34.

Windahl, S. H., Vidal, O., Andersson, G., Gustafsson, J. A., and Ohlsson, C. (1999). Increased cortical bone mineral content but unchanged trabecular bone mineral density in female ERbeta(-/-) mice. J. Clin. Invest. *104*, 895-901.

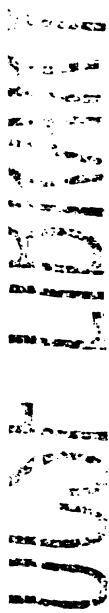


Figure Legends

Figure 1

A. Overall Structure of the THC-ERα LBD-GRIP1 NR Box II Peptide Complex. Two orthogonal views of the THC-ERα LBD-NR box II peptide complex. The coactivator peptide and the LBD are shown in ribbon representation. The peptide is colored gold and helix 12 is colored magenta. Helices 3, 4 and 5 are colored blue and THC, shown in spacefilling representation, is colored green.

B. Overall Structure of the THC-ERβ LBD Complex. Two orthogonal views of the THC-ERβ LBD complex similar to those of the ERα complex in **A.**. The LBD is depicted in ribbon representation. Helix 12 is colored magenta and helices 3, 4 and 5 are colored blue. THC, in spacefilling representation, is colored red

Figure 2 THC Interactions with the ERα LBD (A.) and the ERβ LBD (B.). Residues that interact with THC are drawn at approximately their true positions. The residues that form van der Waals contacts with THC are depicted as labeled arcs with radial spokes that point towards the atoms of THC with which they interact. The residues that hydrogen bond to the ligand are shown in ball-and-stick representation. Hydrogen bonds are represented as dashed cyan lines; the distance of each bond is given. The ligand rings and the individual ligand atoms are labeled.

Figure 3 Interaction of the ERα LBD and the ERβ LBD with the A ring of THC. The structures of the ERα LBD and the ERβ LBD bound to THC were overlapped using the Cα coordinates of residues 306 to 526 from ERα and residues 259 to 476 from ERβ. The THC molecule bound to the ERα LBD is colored green and the THC molecule bound to the ERβ LBD is colored red. The ERα residue numbers are listed above the ERβ residue numbers. Hydrogen bonds are depicted as dashed orange bonds. In the THC-ERα LBD, the phenolic A ring of THC forms van der Waals contacts with the side chains of Ala 350, Leu 387, Leu 391, and Phe 404 and forms hydrogen bonds with the side chain carboxylate of Glu 353, the guanidinium group of Arg 394, and a buried water molecule. In the THC-ERβ LBD, the A ring of THC only packs against the side chains of Phe 356, Ala 302, and Leu 339 and forms a single hydrogen bond with the side chain of Glu 305.

Figure 4 Interaction of the ERα LBD and the ERβ LBD with the A′ ring of THC. The structures of the two complexes were overlapped as in Figure 3. The THC molecule bound to the ERα LBD is colored green and the THC molecule bound to the ERβ LBD is colored red. The ERα and ERβ LBDs are colored gray and blue respectively. The ERα residue numbers are listed above the ERβ residue numbers. Hydrogen bonds are depicted as dashed orange bonds. In the THC-ERα LBD complex, the A′ ring of THC forms van der Waals contacts with Met 421 and Ile 424 and the A′ ring phenolic hydroxyl forms a hydrogen bond with the side chain imidazole of His 524. In the THC-ERβ complex, the A′ ring of

THC does not interact with any of the equivalent residues (Ile 373, Ile 376, and His 475).

Figure 5 The Helix 12/ERβ LBD Interface

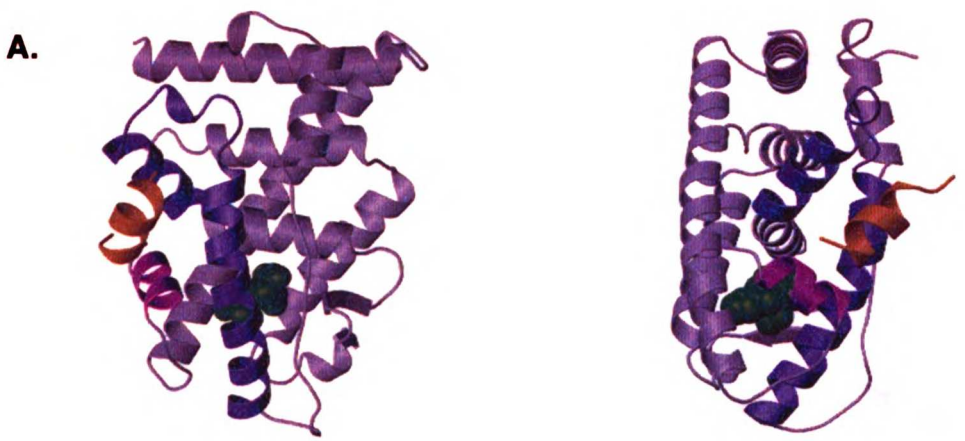
A close-up view of the THC-ERβ LBD complex showing helix 12 bound to the static region of the coactivator binding site (Chapter 1). Residues 481-501 are depicted a Cα worm and residues 487-497 colored magenta. The side chains of Val 487, Leu 491, Leu 494, and Leu 495 are shown. The side chains of only those LBD residues which interact with these residues from helix 12 are shown and are colored by atom type (carbon and sulfur atoms are colored green, oxygen atoms are colored red and nitrogen atoms are colored blue).

Figure 6 Ligands Perturb the Equilibrium Between Different Conformations of Helix 12. Schematic representations of the proposed conformational equilibria of the apo- and the various liganded states of the LBD are depicted. The three conformations of helix 12 (magenta), the GEN/THC-bound conformation, the OHT/RAL-bound conformation, and the E2/DES-bound conformation are labeled "Inactive I", "Inactive II", and "Active", respectively.

Figure 7 Alignment of the Sequences of the ERα LBD and the ERβ LBD. Identical residues in the sequences of the human ERα and ERβ LBDs are boxed. Residues which interact with the ligands are highlighted in green. The only residues that interact with THC which are different in the two ERs are Leu 384/Met 336 and Met 421/Ile 373.

Figure 8 THC Promotes Different Conformations of the ERα and ERβ LBDs. Ribbon representations of the THC-ERα LBD complex (without the coactivator peptide) and THC-ERβ LBD complex. In each complex, THC is shown in spacefilling representation and helix 12 is colored magenta. The regions of the two complexes which differ in secondary structural content are colored red.

Figure 1



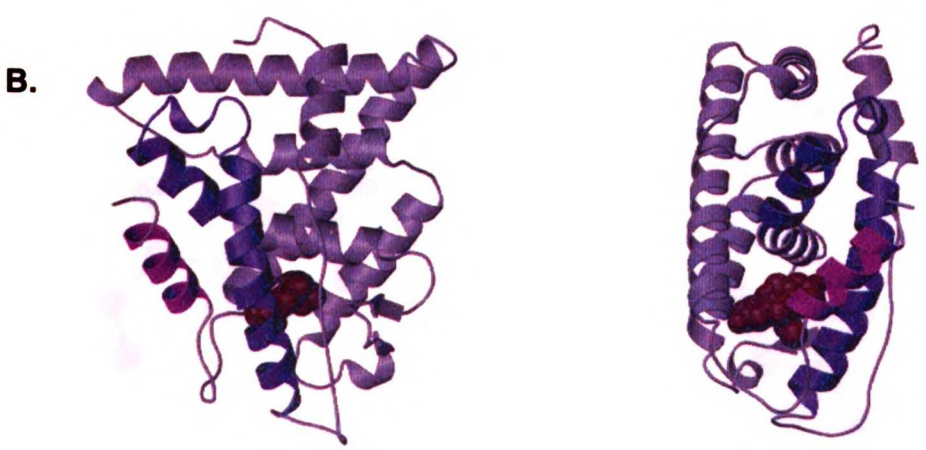


Figure 2A

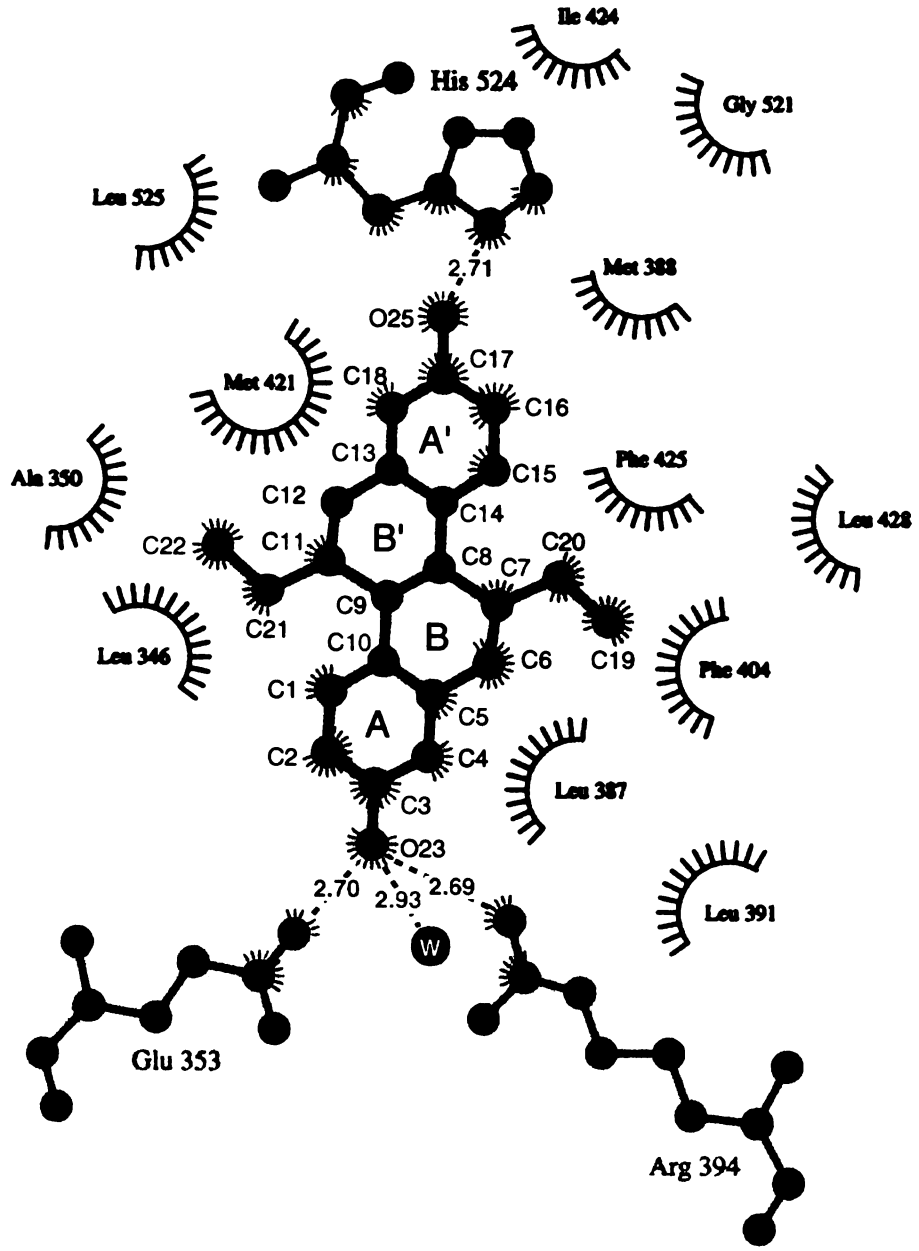

Ile 424
His 524
Gly 521
Leu 525
2.71
Met 388
O25
C17
Met 421
C18
C16
A'
C13
C15
Phe 425
Ala 350
C12
C14
Leu 428
C22
C11
B'
C8
C7
C20
C9
C21
C10
B
Leu 346
C6
C19
Phe 404
C1
C5
A
C2
C4
Leu 387
C3
O23
2.70
2.93
2.69
Leu 391
W
Glu 353
Arg 394

Figure 2B

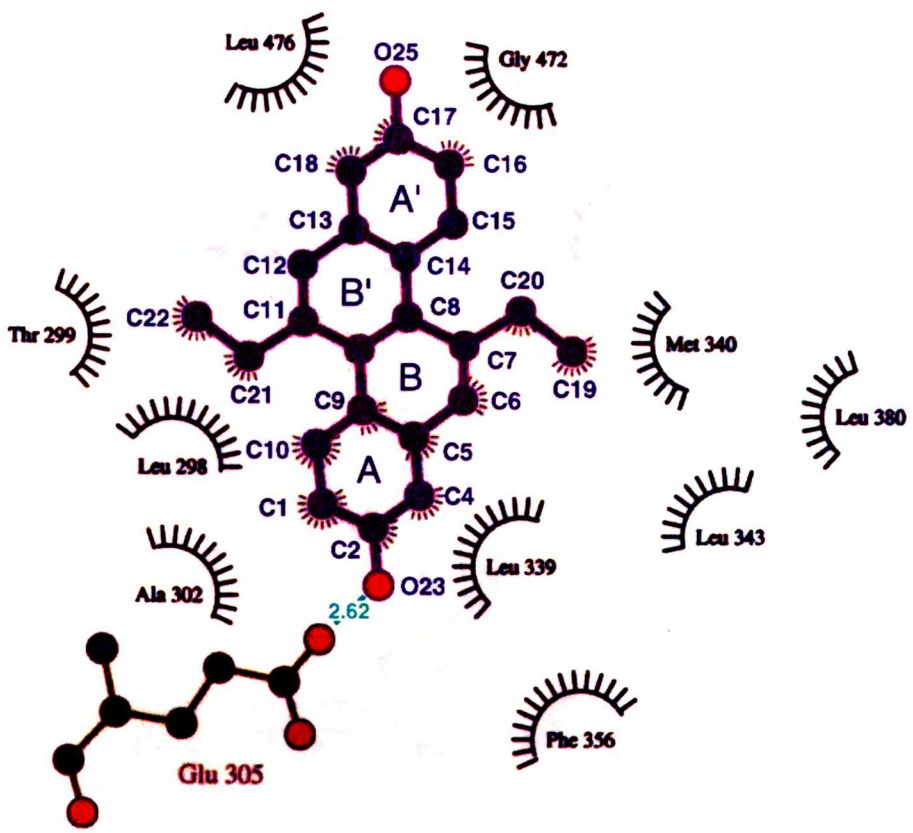


Figure 3

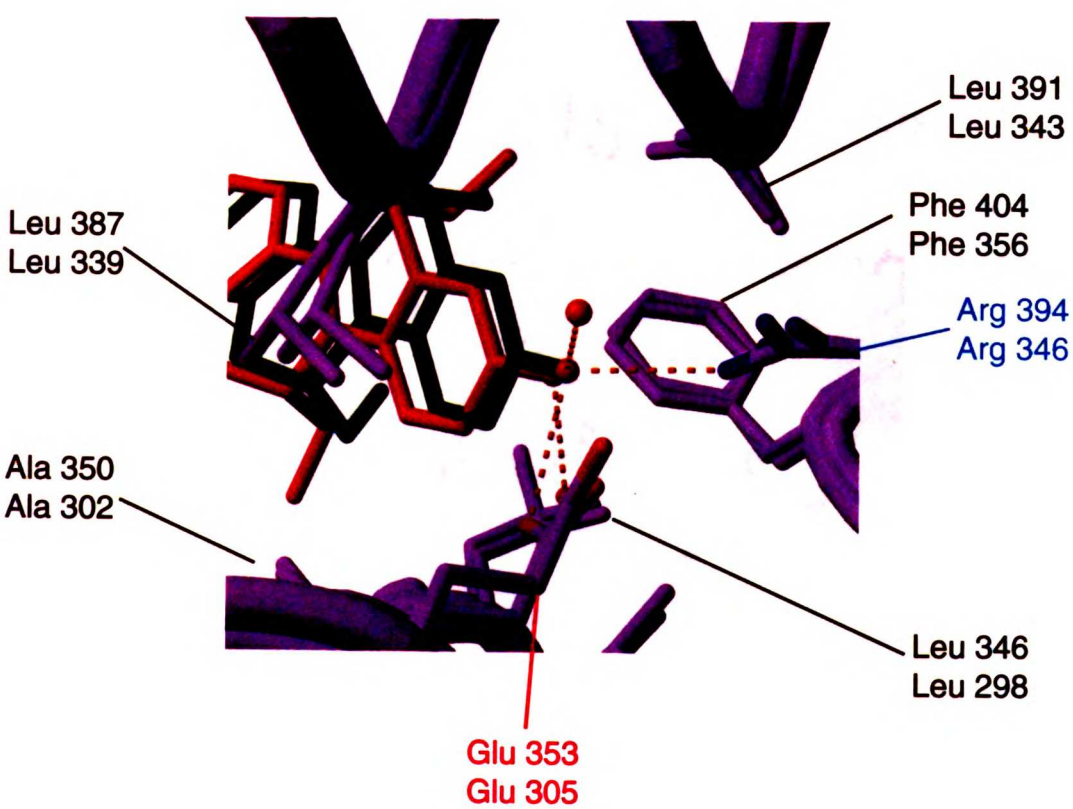


Figure 4

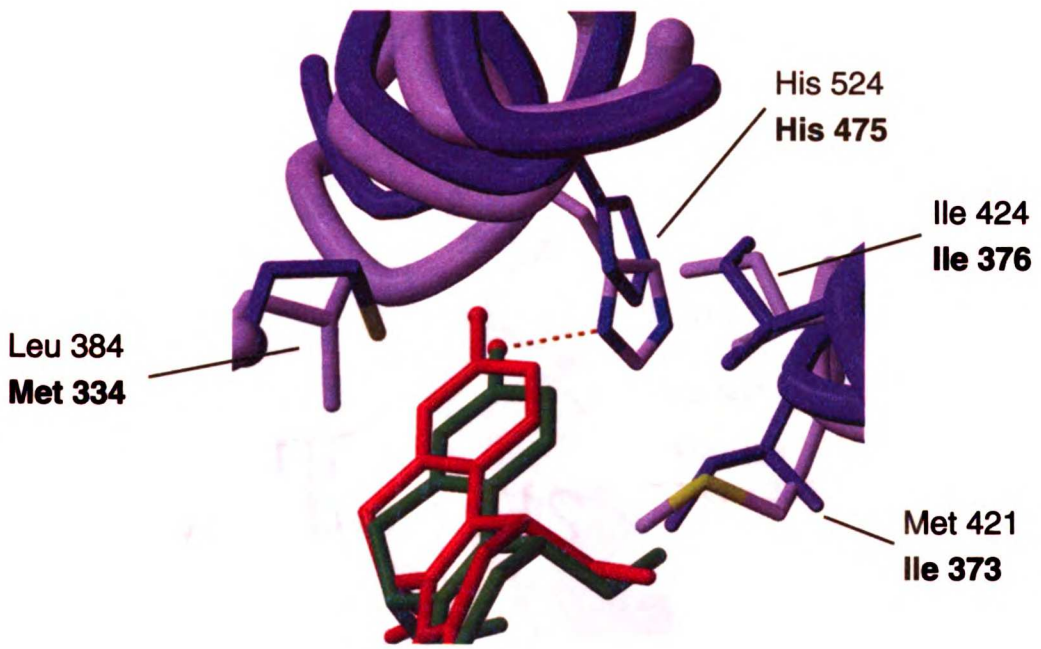


Figure 5

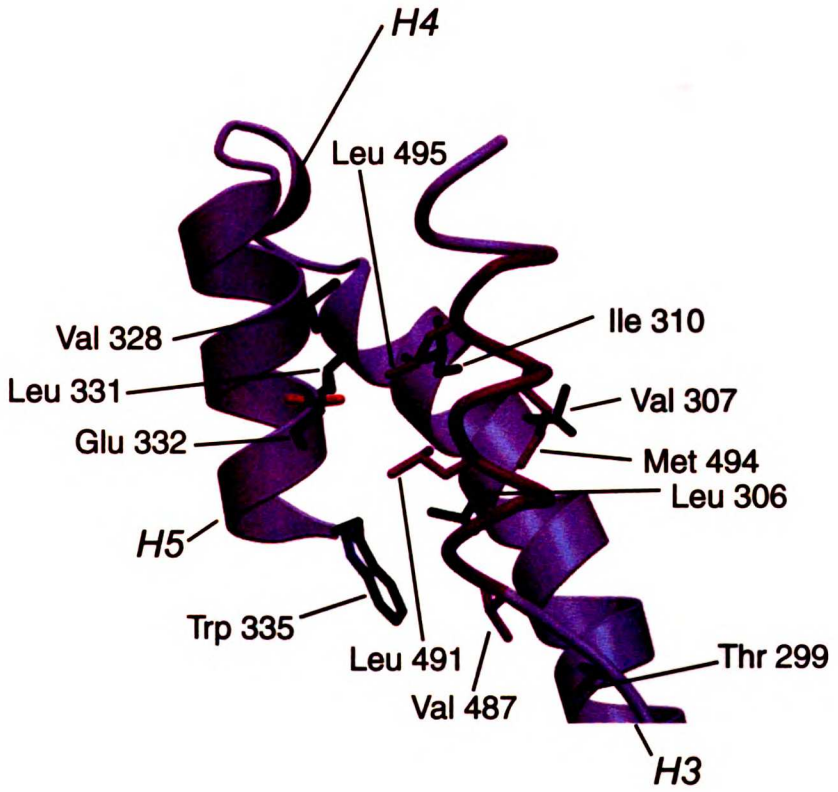
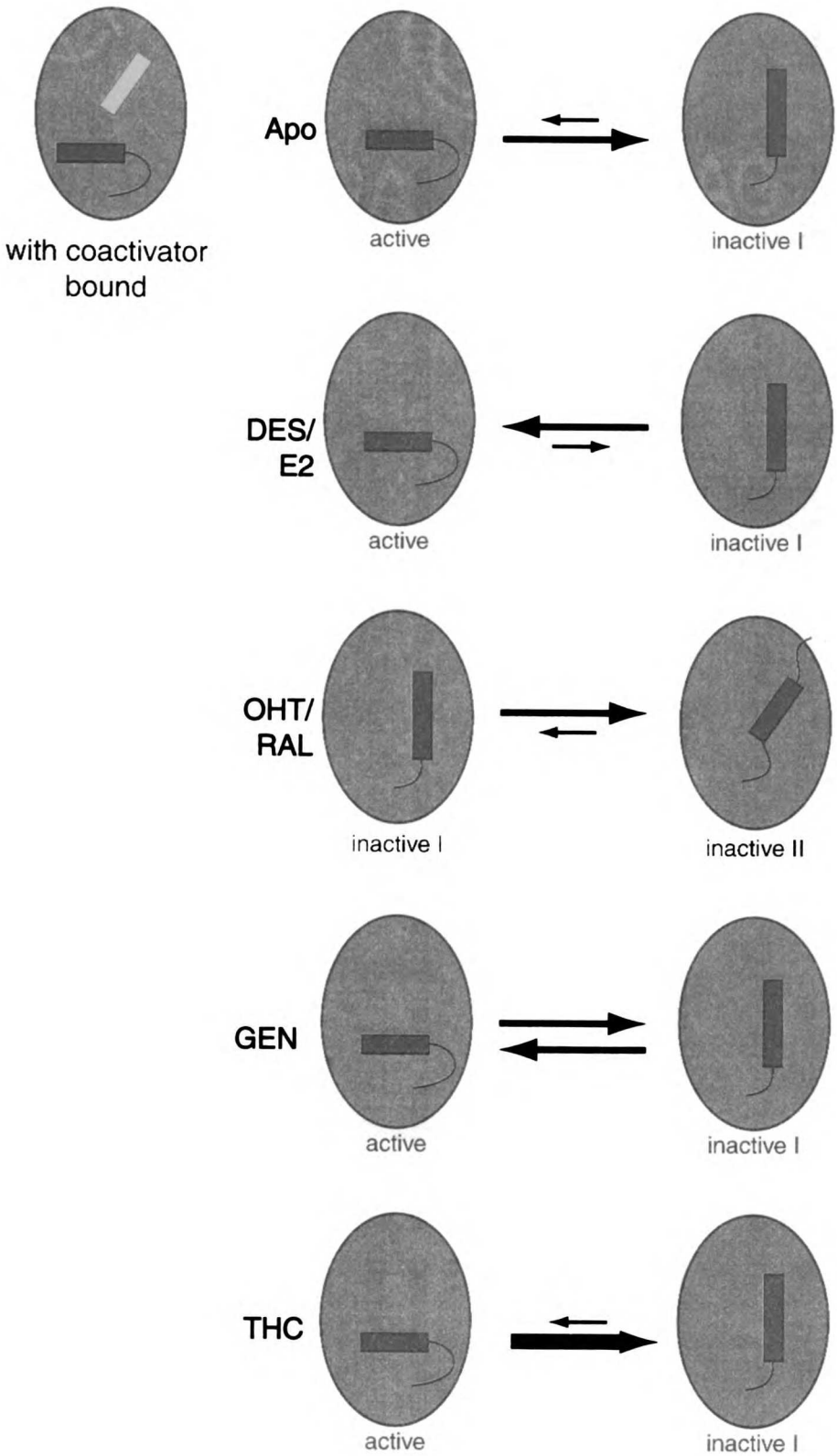

H4
Leu 495
Val 328
Leu 331
Glu 332
H5
Trp 335
Leu 491
Val 487
Ile 310
Val 307
Met 494
Leu 306
Thr 299
H3

Figure 6



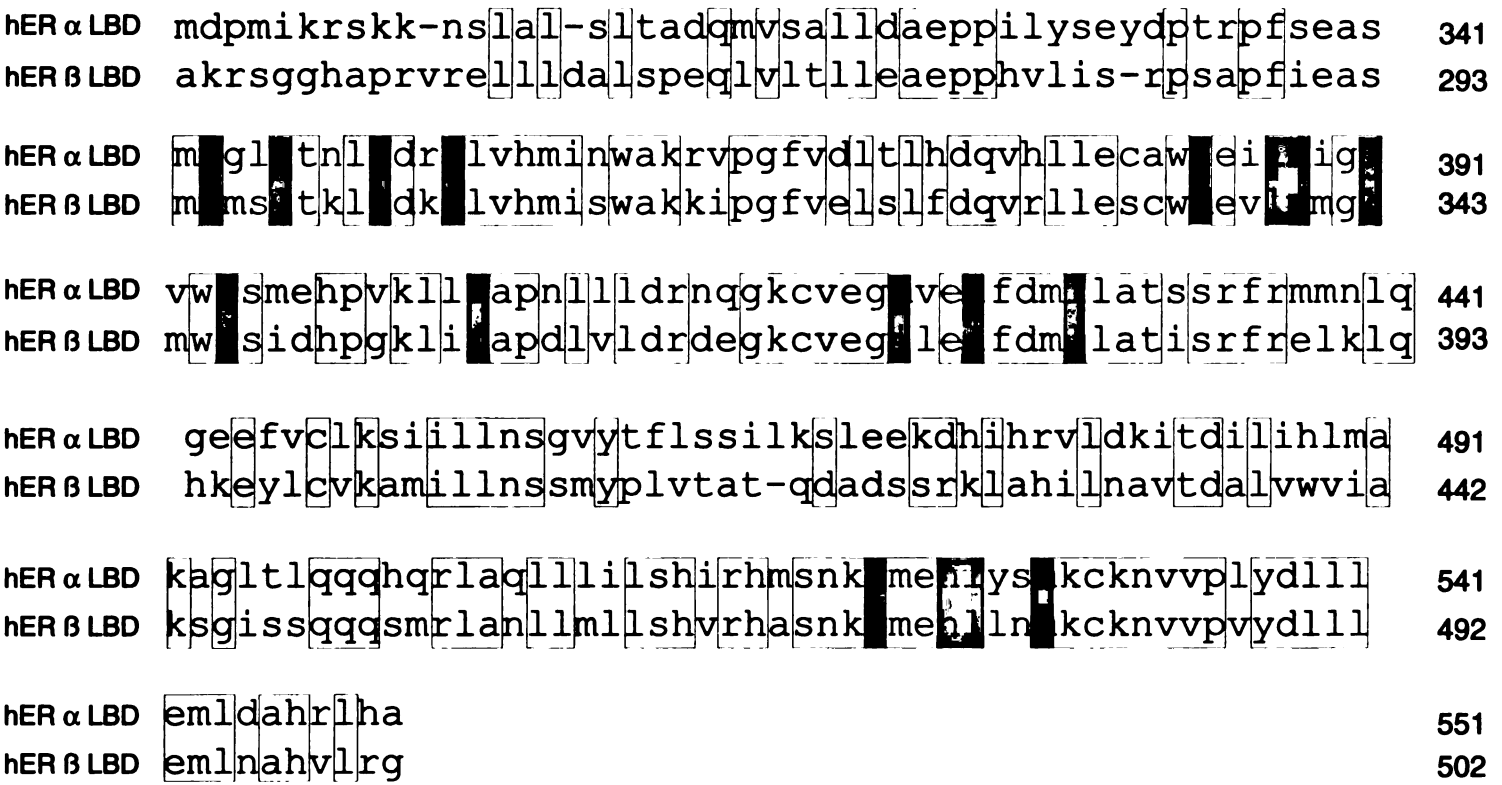

hER α LBD mdpmikrskk-nslal-sltadqmvsalldaeppilyseydptrpfseas 341
hER β LBD akrsgghaprvrellldalspeqlvltlleaepphvlis-rpsapfieas 293
hER α LBD m gl tnl dr lvhminwakrvpgfvdltlhdqvhllecaw ei i ig 391
hER β LBD m ms tkl dk lvhmiswakkipgfvelslfdqvrllescw ev mg 343
hER α LBD vw smehpvkll apnlllldrnqgkcveg ve fdm latssrfrmmnlq 441
hER β LBD mw sidhpgkli apdlvldrdegkcveg le fdm latisrfrelklq 393
hER α LBD geefvclksiillnsgvytflssilksleekdhihrvldkitdilihlma 491
hER β LBD hkeylcvkamillnssmyplvtat-qdadssrklahilnavtdalvwvia 442
hER α LBD kagltlqqqhqrlaqlllilshirhmsnk me ys kcknvvplydlll 541
hER β LBD ksgissqqqsmrlanllmllshvrhasnk me ln kcknvvpvydlll 492
hER α LBD emldahrlha 551
hER β LBD emlnahvlrg 502

Figure 7

Figure 8

Table 1: Summary of Crystallographic Statistics

Data Collection		
Complex	THC-ERα LBD-GRIP1 NR Box II Peptide	THC-ERβ LBD
Space group	$P2_1$	R3
Resolution	1.95	2.75
Observations	119409	78505
Unique	34533	18303
Completeness (%)	97.9	99.7
R_{sym} (%)[a]	6.5	4.9
Average I/σI	12.3	13.5
Refinement		
Resolution	1.95	2.85
Reflections	34533	16521
No. of non-hydrogen atoms	3891	3704
R_{cryst} (%)[b]/R_{free} (%)	24.7/27.1	27.8/31.5
Bond r.m.s. deviation (Å)	0.009	0.011
Angle r.m.s. deviation (°)	1.23	1.42
Average B factor (Å^2)	35.7	40.8

[a] $R_{sym} = \Sigma_i |I_i - \langle I_i \rangle| / \Sigma_i I_i$, where $\langle I_i \rangle$ is the average intensity over symmetry equivalents.

[b] $R_{cryst} = \Sigma |F_o - F_c| / \Sigma |F_o|$.

Chapter 3

Structural Studies of HtpG, the *Escherichia coli* 90 Kilodalton Heat Shock Protein

Andrew K. Shiau and David A. Agard

Howard Hughes Medical Institute

and the Department of Biochemistry and Biophysics

University of California at San Francisco

San Francisco, California 94143

Introduction

The exposure of both prokaryotic and eukaryotic cells to sudden increases in temperature leads to the transient elevated expression of certain proteins called heat shock proteins. Rather than serving as specialized components of a cellular stress response auxiliary to normal cell physiology, many heat shock proteins play a central role in cellular protein biogenesis, under all environmental conditions, by functioning as molecular chaperones (Bukau, 1999; Bukau and Horwich, 1998; Hartl, 1996).

As protein molecules fold or unfold *in vitro*, they transiently expose hydrophobic surfaces which may cause them to aggregate nonspecifically. *In vivo* the problem of these potential nonproductive interactions between partially folded protein molecules is exacerbated, even under normal conditions, by the extremely high concentrations of proteins within the cell. Molecular chaperones limit nonspecific aggregation by interacting with the hydrophobic surfaces exposed by partially folded protein molecules. Given that both the strength of hydrophobic interactions and the probability of folded protein molecules partially or completely unfolding increases with temperature, the problems of aggregating proteins and hence, the demand for molecular chaperones, become particularly acute during heat shock. The increased expression of heat shock proteins is therefore a central component of the cellular response to thermal stress (Bukau, 1999; Bukau and Horwich, 1998; Hartl, 1996).

One of the largest families of heat shock proteins is the 90 kilodalton heat shock protein (hsp90) family (Scheibel and Buchner, 1997). Members of the hsp90 family have been identified in many prokaryotic organisms including Gram-positive bacteria, such as *Bacillus subtilis*, and Gram-negative bacteria, such as *Escherichia coli*, and in virtually all eukaryotic organisms (Gupta, 1995; Schulz et al., 1997). The amino acid sequences of hsp90 proteins have been highly conserved throughout evolution. For example, the sequence of the *Escherichia coli* hsp90, HtpG (high temperature-inducible protein G) is 41% identical to that of the *Drosophila melanogaster* protein, hsp83, and 42% identical to that of the human proteins (Bardwell and Craig, 1987). In addition, hsp90 is one of the most abundant cellular proteins. Under normal conditions, hsp90 proteins comprise 0.5% of total cytosolic protein in *E. coli* (Herendeen et al., 1979) and between 1 to 2% of total cytosolic protein in eukaryotic cells (Borkovich et al., 1989). In mammalian cells, an organelle-specific hsp90 family member, glucose responsive protein 94 (GRP94) is the most abundant protein in the endoplasmic reticulum (Koch et al., 1986).

Consistent with its playing a role in the heat shock response, hsp90 expression increases during heat shock. The levels of HtpG and of hsp83 from *D. melanogaster* increase five-fold and ten- to fifteen-fold respectively, during heat shock (Herendeen et al., 1979; Xiao and Lis, 1989). Many eukaryotic organisms including the yeast, *Saccharomyces cerevisiae*, and various vertebrate organisms, express two cytosolic isoforms of hsp90

(Borkovich et al., 1989; Krone and Sass, 1994). One of these isoforms, hsc83 in *S. cerevisiae* and hsp90β in vertebrates, is expressed at a relatively high levels at normal temperatures and is only weakly induced (two- to three-fold) at high temperatures. The other, hsp82 in *S. cerevisiae* and hsp90α in vertebrates, is expressed at relatively low levels at normal temperatures and can be induced up to fifteen-fold during heat shock (Borkovich et al., 1989; Krone and Sass, 1994).

Genetic studies suggest that hsp90 may be necessary for cellular functions even under normal conditions. Although the HtpG gene is not absolutely required for the viability of *E. coli*, HtpG deletion mutants display a significant growth defect relative to wild-type cells which is exaggerated at high temperatures (Bardwell and Craig, 1988). In other bacteria, such as the cyanobacterium, Synechococcus sp. PCC 7942, the deletion of hsp90 is lethal at all temperatures (Tanaka and Nakamoto, 1999). Similarly, eukaryotic organisms appear to absolutely require hsp90 for viability; deletion of both hsp90 isoforms in *S. cerevisiae* is lethal as is deletion of hsp83 in *D. melanogaster* (Borkovich et al., 1989; Cutforth and Rubin, 1994). Even more remarkably, there is a dramatic increase in the occurrence of morphological changes and developmental defects in strains of *D. melanogaster* that are compromised in hsp83 activity (Rutherford and Lindquist, 1998).

Functionally, members of the hsp90 family, like members of the GroEL/CCT and 70 kilodalton heat shock protein (hsp70) families, act as molecular chaperones (Buchner, 1999). Hsp90 recognizes nonnative folding

intermediates of many types of proteins and prevents the irreversible aggregation of these species during the refolding or unfolding of these proteins *in vitro*. In the case of certain substrates, such as certain Fabs, citrate synthase, and insulin, the addition of hsp90 alone is able to reduce aggregation and improve the yields of active protein (Buchner, 1999). Hsp90 is able to maintain other substrates, such as β-galactosidase, in a nonnative, folding-competent state that can be recognized by other molecular chaperones, such as hsp70, which promote folding to the native state (Freeman et al., 1996).

In vivo, hsp90 interacts with a large number of substrates including many signalling molecules. GRP94 is involved in the maturation process of immunoglobulins (Melnick et al., 1994). Hsp90 binds many serine/threonine and tyrosine kinases including viral $pp60^{src}$ kinase, sevenless, c-raf, cdk4, eIF-2α kinase, casein kinase II, wee1, ste11, and gcn2 (Buchner, 1999). In addition, hsp90 is necessary for the function of the tumor suppressor protein, p53; the transcription factor, heat shock factor-1; an isozyme of nitric oxide synthase; and certain viral reverse transcriptases (Buchner, 1999).

Hsp90, however, has been best-studied as a protein that is necessary for the activity of the steroid hormone receptors (SHRs), including the androgen receptor, the estrogen receptor α (ERα), the estrogen receptor β (ERβ), the glucocorticoid receptor, the mineralocorticoid receptor, and the progesterone receptor (PR) (Bohen and Yamamoto, 1994; Toft, 1999) and other nuclear receptors (Holley and Yamamoto, 1995). Genetic experiments in yeast

provided the first evidence that the interaction between hsp90 and an unliganded or apo-SHR maintains the ability of the SHR to bind ligand (Picard et al., 1990). These results are consistent with the observation that the ligand binding domains (LBDs) of SHRs are necessary and sufficient for hsp90 interaction (Howard et al., 1990). Biochemical experiments in both rabbit reticulocyte lysate and chick oviduct lystate indicate that hsp90 forms the core of a chaperone machine that interacts, through a dynamic process, with apo-SHRs (Toft, 1999). In the first step of the cycle, hsp70 and two hsp70 accessory proteins (hsp40 and hsp70 interacting protein, Hip) associate in an ATP-dependent manner with the SHR. The hsp70 in this complex is then recognized by a protein known as Hop/p60/Sti1 bound to a dimer of hsp90. Through another round of ATP hydrolysis, Hop/p60/Sti1, hsp70, hsp40, and Hip dissociate from this intermediate complex. Concurrently, another chaperone, p23, in the presence of ATP, recognizes the hsp90 bound to the SHR. Finally, various peptidyl-prolyl isomerases, such as FKBP51, FKBP52, and cyclophilin 40, bind the SHR/hsp90/p23 complex and form the final mature complex. The SHR within this complex is now finally able to bind hormone. In the presence of ligand, hsp90 and p23 are released and the SHR translocates to the nucleus where it binds to its cognate response element and either activates or represses transcription. In the absence of ligand, hsp90 and p23 are released and the SHR begins the cycle again by simply rebinding the hsp70/hsp40/Hip complex.

It is perhaps not surprising that SHRs require such an elaborate mechanism for their activity. Ligands form large parts of the cores of the lower subdomains of the LBDs of ERα, ERβ, and PR (Brzozowski et al., 1997; Pike et al., 1999; Tanenbaum et al., 1998; Williams and Sigler, 1998; Chapters 1 and 2). As a result, the LBD of an apo-SHR would be expected to contain a large hydrophobic cavity and, hence, be somewhat unstable. Consistent with this hypothesis, apo-SHR LBDs are prone to proteolysis (Allan et al., 1992; Beekman et al., 1993; Kallio et al., 1994) and are soluble only under denaturing conditions (Seielstad et al., 1995; Williams and Sigler, 1998). SHR LBDs therefore in many ways resemble the folding intermediates that hsp90 has evolved to recognize as a molecular chaperone. In this way, SHRs have coopted the ability of hsp90 to stabilize nonnative species into facilitating the conformational changes they require for ligand-regulated signalling.

X-ray crystallographic and electron microscopic studies of the chaperones GroEL and hsp70 have been instrumental in elucidating their mechanisms of action (Bukau and Horwich, 1998). Structural studies of hsp90, although somewhat less extensive, have also begun to reveal the details of hsp90 function (Prodromou et al., 1997; Prodromou et al., 1997; Stebbins et al., 1997). Limited proteolytic and sequence analysis have indicated that hsp90 has three separable domains (N-terminal, Middle, and C-terminal) (Stebbins et al., 1997). The high resolution x-ray crystal structures of the N-terminal domains of the yeast hsp82 and the human hsp90α, reveal that the domain is an α/β sandwich composed of a highly-twisted, eight-

stranded β-sheet flanked by nine α-helices (Prodromou et al., 1997; Stebbins et al., 1997). Most notably these structures revealed that the hsp90 N-terminal domain shares significant structural homology with the ATP-binding domains of the *E. coli* topoisomerase, DNA Gyrase B, and the *E. coli* DNA repair protein, MutL (Ban et al., 1999; Ban and Yang, 1998; Prodromou et al., 1997). These results inspired a number of structural, biochemical, and genetic studies that demonstrate that the N-terminal domain does indeed bind adenine nucleotides and that full-length hsp90 has a slow, but significant, ATPase activity which is necessary for its essential functions in yeast (Obermann et al., 1998; Panaretou et al., 1998).

Previous biochemical and mutagenesis studies indicate that hsp90 is functional as a dimer (Meng et al., 1996; Minami et al., 1994). In electron micrographs, mammalian hsp90 and GRP94 have been shown form dimers with trinodular structures (Maruya et al., 1999; Wearsch and Nicchitta, 1996). The hsp90α dimer appears to be quite flexible; high temperatures or ATP appear to be able to induce a conformation in which the N-terminal domains of the dimer interact (Maruya et al., 1999). This "circular" structure is apparently consistent with the quaternary structural arrangement of the hsp82 N-terminal domain. The hsp82 N-terminal domain crystallizes as a dimer in which the most C-terminal strands of two monomers interact in an antiparallel manner, allowing the formation of a continuous 16-stranded sheet. A roughly cylindrical channel formed by this sheet between the two

monomers has been proposed to be a binding site for peptide substrates (Prodromou et al., 1997).

Despite these results, the structural mechanisms by which hsp90 recognizes substrates, such as SHRs, and the role of ATP in modulating hsp90 activity remain unclear. In order to begin addressing these issues, I have determined the structure of the *E. coli* hsp90, HtpG to ~4.0 Å. Although the data are not of sufficient resolution for the construction of a complete atomic model of HtpG, they have allowed the preliminary tracing of the backbone of a portion of the Middle domain of HtpG. The Middle domains of HtpG, DNA Gyrase B, and MutL, like their respective N-terminal domans, share extensive structural homology. These results reveal a plausible model for the regulation of hsp90 activity by ATP and the interaction between hsp90 and SHR LBDs.

Results and Discussion

Crystallization

Crystals of full-length, wild-type *Escherichia coli* HtpG were obtained under a range of conditions. Based on morphology, these crystals could be classified as one of three types: 1) twinned plates; 2) twinned prisms; or 3) hexagonal bipyramids or hexagonal rods (Figure 1). Large, thin, twinned plates (up to 1 mm x 0.5 mm x 0.01 mm) were obtained in the absence or presence of various nucleotides (including ADP, ATPγS, and AMPPNP) using ammonium sulfate, lithium sulfate, sodium sulfate, or ammonium phosphate as precipitants (Figure 1). Diffraction was not observed from these crystals. Small, twinned prisms (up to 200 μm x 50 μm x 50 μm) were obtained in the presence of the anti-tumor agent, geldanamycin (Whitesell et al., 1994), using ammonium sulfate as the precipitant (Figure 1). At room temperature, these crystals diffracted x-rays to ~10 Å using a conventional source. Hexagonal rods or bipyramids were obtained using sodium formate, sodium citrate, or polyethylene glycol (PEG) 3000 in the presence of calcium chloride, barium chloride, magnesium chloride, sodium acetate, or lithium sulfate as precipitants. At room temperature, these crystals diffracted x-rays to between 5 to 6 Å at synchrotron sources. Screens for additives revealed that the inclusion of certain nonionic detergents in the PEG 3000/magnesium chloride precipitant combination yielded extremely large crystals (up to 0.5 mm x 0.5 mm x 1.0 mm) (Figure 1) that diffracted x-rays anisotropically to ~6

Å using conventional sources (Figure 2) and ~3.5 Å at synchrotron sources (Figure 3). Of the ~30 detergents tested, only β-octylglucoside (but only certain batches obtained from the company Anatrace) produced reproducible results. Screens for additional additives that improved the diffraction limit of these crystals were unsuccessful.

Data Collection

The crystals of HtpG obtained from the PEG 3000/magnesium chloride/detergent precipitant combination are problematic for several reasons. They are fragile and radiation sensitive, diffract x-rays relatively weakly and anisotropically (to ~3.2 Å along a* and b* and to ~4 Å along c*) (Figure 3), and have one extremely long unit cell axis (c ~500 Å). Each of these problems had to be solved before data could be measured from these crystals.

Protein crystals are typically manipulated either by pipetting them into thin-walled glass or quartz capillaries (Rayment, 1985) or by suspending them in thin films of mother liquor formed in synthetic fiber loops (Rodgers, 1997). Use of either of these techniques in the manipulation of crystals of HtpG frequently damaged these crystals. HtpG crystals would often crack or shatter as they were being drawn into capillaries, presumably due the shearing forces of the suction. Loops made from nylon, rayon or glass tended to bow or twist, causing the crystals to crack during manipulation. As a result, all transfers were performed using loops (1-3 mm in diameter) made from small gauge (0.002 to 0.005 inch thick) metal wire (platinum or Nichrome). Unlike their

synthetic fiber or glass counterparts, metal loops did not bow or twist. Metal loops of this size could also support extremely thick films of mother liquor (~1-3 μL) which helped to minimize the mechanical stress exerted on the crystals during transfers.

At room temperature, the quality of the diffraction from HtpG crystals rapidly degraded during data collection. The discovery of a useful method of manipulating these crystals facilitated the development of a procedure to flash cool them and prevent this radiation damage. After screening various procedures and cryoprotectant solutions (Rodgers, 1997), it was found that the crystals could only be succesfully flash frozen after they had been introduced, rapidly and stepwise, into a cryoprotectant solution containing ethylene glycol (Experimental Procedures). The design of the transfer protocol was guided by three observations. First, cryoprotectant solutions containing many organic solvents (other than ethylene glycol), such as glycerol and 2-methyl-2,4-pentanediol, or sugars, such as glucose or sucrose, generally compromised the quality of the diffraction from these crystals. Second, because prolonged exposure (> 10 minutes) of the crystals to the ethylene glycol-based cryoprotectant generally led to loss of diffraction, the transfer procedure had to be executed as quickly as possible. Third, if the crystals were directly transferred into the ethylene glycol-based cryoprotectant, the crystals would either crack or shatter (perhaps due to osmotic stress).

The actual techniques used to mount and freeze the crystals (after they had been introduced into the cryoprotectant) were also critical to the success

of the flash cooling process. Crystals of HtpG could only be successfully frozen in films of cryoprotectant that were no thicker than the gauge of the metal wire used to make the loops. Therefore, after suspending the crystals in films of the cryoprotectant formed in metal wire loops, cryoprotectant was gently wicked away with small pieces of Kimwipe or filter paper until the films had the appropriate thickness. In most cases, the crystals were frozen by plunging them into a bath of liquid nitrogen. This freezing technique (as opposed to freezing the crystals directly in the cold nitrogen stream of a cryostat) resulted in greatest isomorphism between frozen crystals. Due to the number of manipulations involved, only one out of four crystals survived the entire flash cooling process.

Diffraction data were measured either at the Stanford Synchrotron Radiation Laboratory (SSRL) beamline 1-5 using Fuji imaging plates or at the Cornell High Energy Synchrotron Source (CHESS) station F1 using a CCD camera. The long crystal-to-detector distances (> 400 mm) required to resolve the reflections along the long axis of these crystals, c*, in combination with the weak diffraction from these crystals precluded data collection at all conventional x-ray sources and many synchrotron sources. The finely focussed, intense beams and the camera geometries at these facilities made data collection feasible.

Because of the size limitations of the active area of the detectors at these beamlines and the long crystal-to-detector distances necessary for data collection, the detectors were offset in the horizontal and vertical directions

prior to data collection. This detector positioning allowed the measurement of only a single quadrant of data at any one time. Therefore, prior to freezing, the crystals were aligned visually in the loops in which they were suspended such that c* (which corresponded to the longest physical dimension of the crystals) was within 5 to 15 degrees of the spindle axis. Because this alignment positioned the shorter (~100 Å) cell axes approximately perpendicular to the rotation axis, this strategy allowed for relatively wide oscillations to be collected on each image. In addition, this crystal alignment minimized the "missing cone" problem and greatly increased the probability of achieving complete data coverage. Between 60 to 100 degrees of data were measured in order to obtain data sets that were between two- and six-fold redundant.

Ultimately, native data were measured using three crystals. Data were measured at SSRL beamline 1-5 and CHESS station F-1 from two crystals grown at ~pH 8. Individually, these data sets were incomplete but when merged, they provided a data set that was ~90 % complete to ~4.1 Å. The other data set (~90% complete to 3.7 Å) was measured at SSRL beamline 1-5 from a crystal grown at ~pH 9. Attempts to scale the pH 8 and the pH 9 data sets indicated a lack of isomorphism between the two data sets.

Heavy Atom Derivatives

All attempts at generating heavy atom derivatives by soaking these crystals in heavy atom solutions were unsuccessful. In most cases, treatment with heavy atom compounds either visibly damaged the crystals or greatly reduced

their ability to diffract x-rays. As a result, alternative approaches to obtaining heavy atom derivatives were taken.

The sequence of wild-type HtpG does not contain any cysteines. A now commonly used strategy for obtaining heavy atom derivatives of crystals of proteins which are naturally devoid of cysteines involves the systematic introduction of cysteines into these proteins by site-directed mutagenesis (Hatfull et al., 1989; Sun et al., 1987). In principle, these engineered cysteines could serve as potential sites of attachment for various mercurial compounds. Serines are frequently selected for this type of mutagenesis because serine is isosteric with cysteine and serines are frequently located on the surfaces of proteins.

Seventeen residues throughout the sequence of HtpG were targeted for mutagenesis (Table 1). Fourteen of these residues (Ser 40, Ser 79, Ser 104, Ser 156, Ser 198, Ser 282, Ser 335, Ser 421, Ser 437, Ser 461, Ser 484, Ser 500, Ser 547, and Ser 590) were serines that were flanked by one charged residue and one hydrophobic residue in the primary sequence of the protein. By this criterion, these residues were predicted to be partially buried on the surface of the protein. The use of this selection criterion was dictated by two competing factors. In order for the planned subsequent modification with mercurial compounds to be successful, these residues would have to be located at the surface of the protein. However, any mobility of attached heavy atoms would dampen the contribution of these heavy atoms to the total scattering of a protein crystal. Therefore, the residues targeted for modification should be

the least mobile, solvent-accessible sites available and hence partially buried. The other three residues selected for mutagenesis were Thr 220, Ala 310, and Val 356. These residues were chosen because the residues in the analogous positions in the sequences of the protein from other species were frequently cysteines.

Of the seventeen mutagenesis reactions performed, fifteen were successful (Table 1). In addition, using the appropriate mutant clones, a mutant was generated with three of the previously described sequence changes (Ser 79 -> Cys, Ser 156 -> Cys, Ala 310 -> Cys). Twelve out of the resulting sixteen mutant proteins, including the variant with three mutations, were soluble and expressed at or near wild-type levels (100 to 150 mg/L of culture) (Figure 4A). Prior to the purification and crystallization of these mutant proteins, a simple and rapid biochemical assay was used to determine if the cysteines of these mutant proteins were indeed solvent accessible. Samples of extracts containing the mutant proteins were treated, in the presence and absence of denaturant, with a derivative of a fluorescent dye that reacts preferentially and irreversibly with free sulfhydryls, Lucifer Yellow Iodoacetamide (LYI). After quenching with excess DTT, the reaction products were separated by SDS-PAGE and visualized both by UV transillumination and by staining with Coomassie blue. In the denatured state, all of the soluble mutant proteins, but not the wild-type protein, were labeled rapidly by LYI (Figure 4B). In the native state, most of the mutant proteins were labeled more slowly by LYI (Figures 4C and 4D). In addition,

there was a large variation in the rates of labeling among the various mutants. The Ser 437 -> Cys protein, which was completely labeled within an hour (assuming that the labeling reactions of the denatured proteins had gone to completion), was the most rapidly labeled and the Ser 282 -> Cys protein, which was incompletely labeled even after 20 hours, was the most slowly labeled (Figures 4C and 4D). These results indicated that cysteines in all of the mutant proteins were solvent accessible (to some degree) and that these cysteines were located in different chemical environments on the surface of the protein.

Mercury can be incorporated into protein crystals either by treating the proteins with the appropriate mercurial compounds prior to crystallization or by treating the protein crystals themselves with mercurial compounds. As a result, crystallization trials of eleven of the cysteine mutants were attempted using either protein that had been treated with methyl mercury nitrate (MMN) (Experimental Procedures) or protein that had been left untreated. Seven mutants (the Ser 79 -> Cys, Ser 156 -> Cys, Ser 282 -> Cys, Ala 310 -> Cys, Ser 484 -> Cys, Ser 590 -> Cys, and Ser 79 -> Cys/ Ser 156 -> Cys/Ala 310 -> Cys variants) yielded usable crystals when pretreated with MMN and eight mutants (the Ser 79 -> Cys, Ser 104 -> Cys, Ser 156 -> Cys, Ser 198 -> Cys, Ser 282 -> Cys, Ala 310 -> Cys, Ser 590 -> Cys, and Ser 79 -> Cys/ Ser 156 -> Cys/Ala 310 -> Cys variants) yielded usable crystals when left untreated.

Data were measured from crystals, grown at pH 8, of the Ser 79 -> Cys, Ser 156 -> Cys, Ser 282 -> Cys, Ala 310 -> Cys, Ser 484 -> Cys, Ser 590 -> Cys, and

Ser 79 -> Cys/ Ser 156 -> Cys/Ala 310 -> Cys variant proteins pretreated with MMN. Data were also measured from crystals, grown at pH 9, of the Ser 79 -> Cys, Ser 156 -> Cys, and Ser 79 -> Cys/ Ser 156 -> Cys/Ala 310 -> Cys variant proteins pretreated with MMN. All crystals of the Ser 198 -> Cys mutant cracked or shattered when treated with a variety of mercurial compounds including MMN. Data were measured from a single crystal, grown at pH 8, of the Ser 104 -> Cys mutant that had been soaked for several hours in 3 mM MMN.

A variant of HtpG which had been labeled biosynthetically with selenomethionine was also made in order to obtain a heavy atom derivative. Selenomethionyl HtpG was expressed in the methionine auxotroph, B834(DE3). Initially, the B834(DE3) cells were grown in LeMaster media (LeMaster and Richards, 1985) supplemented with 1 X Kao and Michayluk Vitamin solution. Unfortunately, it was found that this strain could grow under these conditions whether the media had been supplemented with methionine or not! This conditional phenotype would prevent the desired incorporation of selenomethionine into the expressed HtpG.

Experimentation with the media revealed that this apparent loss of methionine auxotrophy was a result of the 1 X Kao and Michayluk Vitamin solution; the strain was found to be auxotrophic for methionine when grown in LeMaster media without vitamin supplementation or in LeMaster media supplemented with thiamine. Analysis of the methionine biosynthetic

pathway provides a simple explanation for the vitamin-dependent prototrophy exhibited by B834 cells.

The last step in methionine biosynthesis involves the methylation of homocysteine. In *E. coli*, this reaction can be carried out by the products of two genes, *metE* and *metH*. *MetE* encodes a vitamin B_{12}-independent transmethylase and *metH* encodes a vitamin B_{12}-dependent transmethylase (Cohen and Saint-Girons, 1987). Because *E. coli* does not synthesize vitamin B_{12}, strains with a mutation in *metE* require either methionine or vitamin B_{12} for growth. Given that Kao and Michayluk Vitamin solution contains vitamin B_{12}, B834(DE3) most likely contain a mutation in the *metE* gene. This observation may explain various reports in the literature of substoichiometric incorporation of selenium in proteins expressed in B834(DE3) (Cohen and Saint-Girons, 1987).

Crystals of selenomethionyl HtpG, which were only grown at pH 8, were even more difficult to handle than those of the wild-type or of any of the mutant proteins. These crystals, which were even more fragile than crystals of the wild-type protein, grew from drops containing thick masses of gummy precipitate. After several hours of manipulation, a small crystal was successfully extracted from one of the drops and mounted. All attempts to mount larger crystals failed.

Phasing and Multicrystal Form Averaging

All of the potential derivative data sets, except the one collected from the crystal of the Ser 104 -> Cys protein, proved useful for phasing. The two native data sets (pH 8 and pH 9) were initially phased separately using data measured from the potential heavy atom derivative crystals grown at the corresponding pHs.

Ficoll density measurements indicated that the asymmetric unit of the crystals of HtpG most likely contained two monomers of HtpG (Westbrook, 1985). The positions of the two mercuries in the asymmetric unit of the MMN-derivatized Ser 484 -> Cys crystals were determined by manual inspection of difference Patterson maps. The positions of the mercuries in the crystals of the other mutants grown at pH 8 and of 18 seleniums (out of a possible 26) in the crystals of the selenomethionyl protein were determined using phases derived these manually-determined mercury positions and difference Fourier maps.

As described above, the pH 8 and pH 9 native data sets were not completely isomorphous; however, they were closely related enough such that the positions of the mercuries in the MMN-derivatized Ser 79 -> Cys, Ser 156 -> Cys, and Ser 79 -> Cys/ Ser 156 -> Cys/Ala 310 -> Cys mutant crystals grown at pH 9 could be initially be approximated using the coordinates determined from the data measured from crystals of the same mutants grown at pH 8. In both cases, in first round of refinement, the B-factors and

occupancies of the heavy atoms were held fixed; only their positions were refined (Tables 2 and 3).

The initial MIR maps, particularly in the pH 8, showed a very clear boundary between solvent and protein regions of the asymmetric unit. However, the density was not readily interpretable with respect to secondary structural elements (as might be expected in maps calculated to ~4 Å resolution). As a result, a real space density averaging scheme was implemented to take advantage of the noncrystallographic symmetry (NCS) present in these crystals and the lack of isomorphism between the pH 8 and pH 9 cells. Using the pH 8 MIR phased map, a mask was drawn around what was interpreted to be one of the monomers in the asymmetric unit. Noncrystallographic symmetry (NCS) operators and intercrystal operators (between the monomers in the pH 8 and pH 9 cells) were derived from the heavy atom positions within the two cells.

Initially, the density for the two monomers within the pH 8 cell was averaged. This averaged density was then projected into the pH 9 cell. In subsequent rounds, the resulting density of all four monomers was then averaged using data to 4.1 Å. The resulting phases were then used to rerefine the heavy atom parameters in both cells. These improved MIR phases were then used in a second round of multicrystal form averaging.

For the second round of averaging, the amplitudes of both native data sets were modified. The average intensities of both data sets fall off very sharply as a function of resolution; the Wilson B-factor of both data sets is

>150 Å^2. In addition, both data sets are also anisotropic with the intensities falling off much more rapidly in the c* direction than in the other two directions. As a result, the high resolution terms (particularly those along c*) of both data sets contribute very little to the resulting maps and, hence, poorly constrain the averaging process. Therefore, the amplitudes and sigmas of both data sets were adjusted such that the overall Wilson B-factor was ~30 Å^2 by multiplication with a positionally-dependent gaussian temperature factor of the form $b_{11}h^2 + b_{12}hk + b_{13}hl + b_{22}k^2 + b_{23}kl + b_{33}l^2$ (Blundell and Johnson, 1976). These adjusted amplitudes were used in a second round of multicrystal form averaging to 4.1 Å. The phasing of the pH 9 data was then extended to 3.7 Å (in 0.01 Å steps) by NCS averaging starting with phases derived from those generated by multicrystal form averaging and the MIR phases (calculated from the rerefined heavy atom parameters).

Monomer Structure

The resulting map of the pH 9 cell was exceptionally clear and many secondary structural elements, including two large β-sheets and over a dozen α-helices were readily identifiable (Figure 5). Consistent with previous biophysical characterization (Spence and Georgopoulos, 1989), HtpG appears to be an elongated, comma-shaped molecule with an overall head-to-tail length of ~100 Å (Figure 5). Maps were intially calculated in both possible enantiomorphic spacegroups, $P6_522$ and $P6_122$. A model of the yeast hsp82 N-terminal domain (Prodromou et al., 1997) could be manually positioned in

the density of a map calculated in $P6_522$ allowing the absolute assignment of the spacegroup.

The HtpG N-terminal domain is an α/β sandwich formed by an eight stranded β-sheet flanked by nine α-helices (Figures 5 and 6). The model of the N-terminal domain is consistent with the predicted positions of the mercuries derived from the derivatized Ser 79 -> Cys and Ser 156 -> Cys mutant data and the presence of two selenium sites corresponding to Met 20 and Met 85. Overall, the conformation of HtpG N-terminal domain appears to resemble those of the N-terminal domains of the yeast and human proteins (Prodromou et al., 1997; Stebbins et al., 1997). The only regions that differ radically in conformation in the structures of HtpG and hsp82 are located near the ATP binding sites of the two proteins. In the structures of the isolated yeast N-terminal domain, residues 94 to 124 (equivalent to residues 95 to 127 in HtpG) form a short helix, a loop, followed by a short 3_{10} helix. This structural segment interacts with a helix and loop formed by residues 10 to 27 (equivalent to residues 11 to 28 in HtpG) and extends away from the ATP binding site (Prodromou et al., 1997). As described above, the N-terminal domains of hsp82, hsp90α, DNA Gyrase B, and MutL share significant structural and functional homology as ATP binding domains. Although somewhat different in terms of secondary structural content, the regions in the N-terminal domains of Gyrase B and MutL equivalent to residues 10 to 27 and 94 to 124 of hsp82 fold over their respective ATP binding sites, sealing the active sites of these proteins (Ban et al., 1999; Wigley et al., 1991). The

conformations of the active site lids of Gyrase B and MutL are thought to be required for the ATPase activity of these proteins (Ban et al., 1999; Wigley et al., 1991). Although the crystals of HtpG were grown in the absence of nucleotide, residues 95 to 127 are folded over the active site (Figure 6). In addition, the N-terminal region of the first helix in HtpG (formed by residues 11 to 28) is pivoted relative to the position of the equivalent helix of hsp82 by over ~15 Å (Figure 6). This position of the N-terminal helix of HtpG prevents the N-terminal residues (residues 4 to 10) from forming one of the strands of the central beta sheet (Figure 6) as do the equivalent residues in the structure of the hsp82 N-terminal domain. If the conformations of the various components of the active site lid in the HtpG structure are required for the ATPase activity of HtpG, this would provide one reason why, in contrast to the full-length hsp82, the isolated N-terminal domains of the yeast protein is only capable of binding ATP and not hydrolyzing it (Prodromou et al., 1997).

The regions of main chain connecting the N-terminal domains and the Middle domains of HtpG and hsp82 also appear to adopt different conformations in the two structures (Figures 5 and 6). In the structure of the hsp82 N-terminal domain, this region of the main chain forms an extension of the most C-terminal strand of the domain. The C-terminal strands of two monomers interact to complete the continuous sixteen-stranded beta sheet seen in the crystallographic dimer (Prodromou et al., 1997). In the structure of HtpG, this connecting region forms a loop and an additional strand which follows the body of the N-terminal domain to connect with the Middle

domain (Figure 5). This observation and the fact that the human hsp90α N-terminal domain has not been observed to form the same crystallographic dimers suggest that the hsp82 N-terminal dimer may be a crystal packing artifact.

The density maps were sufficient quality to allow a partial trace of the Middle domain of HtpG (Figures 5 and 7). This trace is consistent with the positions of the mercury positions deduced from the derivatized Ser 282 -> Cys and the Ala 310 -> Cys data and the presence of several selenium markers. This portion of the HtpG Middle domain is formed by a five stranded β-sheet flanked by two α-helices (Figure 7). At least part of this domain, like the N-terminal domain, appears to be of the mixed α/β type and shares a similar topology with the equivalent regions of Gyrase B and MutL (Ban et al., 1999; Wigley et al., 1991). The only difference in the topology between this portion of the HtpG Middle domain and those of Gyrase B and MutL is the absence of the crossover helix between the fourth and fifth strands of the HtpG beta sheet (Figure 7). The Middle domains of Gyrase B and MutL contribute the critical active site residues which interact with the γ-phosphate of the nucleotide (Ban et al., 1999; Wigley et al., 1991). If the structural homology between HtpG, Gyrase B, and MutL translates into functional homology, this would provide a second reason why the isolated N-terminal domains of hsp82 and hsp90β are incapable of hydrolyzing ATP.

Unlike the density for the other two domains, very little of the density for the C-terminal domain could be interpreted (Figure 5). The most notable feature of the domain that could be identified (with the aid of mercury positions derived from the derivatized Ser 590 -> Cys mutant data) was that the C-terminus of the molecule appears to consist of two long α-helices joined by a short loop (Figure 5).

Dimer Structure

Consistent with previous biochemical and electron microscopy studies of GRP94 and human hsp90α (Maruya et al., 1999; Wearsch and Nicchitta, 1996), the HtpG dimer appears to formed by two molecules arranged in an antiparallel arrangement. The two molecules form the shape of the letter "C" being ~140 Å at the widest point and ~90 Å tall (Figure 8) in agreement with the equivalent dimensions of the GRP94 and human hsp90α dimers (Maruya et al., 1999; Wearsch and Nicchitta, 1996). When viewed appropriately, the HtpG dimer appears to possess the trinodular structure of GRP94 and hsp90α observed in electron micrographs (Maruya et al., 1999; Wearsch and Nicchitta, 1996). The dimer interface appears to be formed solely by the interaction of the most C-terminal helices of the two monomers, which together form a four-helix bundle (Figure 8). This is consistent with the results of mutagenesis studies of human hsp90α which showed that deletion of the most C-terminal 49 residues of the protein disrupts dimer formation

(Minami et al., 1994). These observations were also confirmed by characterizing the dimerization properties of two fragments of HtpG, a C-terminal truncation mutant ending at residue 517 (N517) and an N-terminal truncation mutant beginning at residue 497 (497C). Both of these domains were expressed recombinantly, purified, and characterized by size exclusion chromatography and glutaraldehyde crosslinking (Figures 9A and 9B). By size exclusion chromatography, the estimated native molecular weights of N517 and 497C are 65 kD, and 40 kD respectively. This indicates that N517 (predicted molecular weight ~60 kD) is a monomer and 497C (predicted molecular weight ~20 kD) forms a dimer. These results were confirmed by crosslinking experiments (Figures 9A and 9B). Whereas the incubation of both the wild-type protein and 497C with glutaraldehyde led to the formation of dimeric crosslinking products, the incubation of N517 under the identical conditions did not. By these assays, in the absence of ATP, the C-terminal ~100 residues of HtpG are necessary and sufficient for the dimerization of the protein.

Like HtpG and other hsp90s, Gyrase B and MutL form constitutive dimers through the interaction of domains C-terminal of their Middle domains (Ban et al., 1999; Wigley et al., 1991). Recent experiments have revealed that both Gyrase B and MutL form additional dimeric interfaces in the presence of nucleotide. Both of these proteins must undergo a conformational change that allows the interaction of the N-terminal domains of these proteins in order for ATP hydrolysis to occur (Ban et al., 1999; Wigley

et al., 1991). The mode of interaction between the N-terminal domains of these proteins is very different from that observed for the hsp82 N-terminal domain described previously. The regions that form the active sites of the N-terminal domains of two monomers interact to form an extensive dimer interface (Figure 8). In this configuration, the first ~15 residues of each N-terminal domain associates with and stabilizes the ATP-bound conformation of the opposing monomer in the dimer (Figure 8) (Ban et al., 1999; Wigley et al., 1991). Given the extensive structural similarity between these proteins and HtpG and other hsp90s, it is tempting to speculate that a similar structural rearrangement may be required for HtpG to undergo ATP hydrolysis. This would be consistent with the observation in electron micrographs of closed ring-like structures in samples of hsp90 in the presence of ATP (Maruya et al., 1999).

This proposed conformational change has interesting implications for the interaction of HtpG with substrates. The central cleft formed between the two monomers of the HtpG dimer is large enough to accommodate a protein the size of an SHR LBD (Figure 10). In contrast, the cleft formed between the two monomers of MutL in its ATP-induced dimer is substantially smaller (Figure 8) (Ban et al., 1999). If HtpG and other hsp90s undergo the same ATP-induced structural transition, the resulting cleft formed by the two monomers of the dimer would only be large enough to interact with either isolated loops or secondary structural elements (such as helices) of substrate proteins. This is consistent with chemical modification and limited proteolysis studies of

the GR LBD, which indicate that the GR LBD (and hence possibly other SHR LBDs) are partially unfolded during the course of hsp90 binding (Modarress et al., 1997; Simons et al., 1989). This partial unfolding event occurs concomitantly with the acquisition of the of the LBD to bind hormone (Modarress et al., 1997; Simons et al., 1989). Given what is known about the structure of SHR LBDs, what type of unfolding event might be linked to the process of ligand binding?

As discussed in Chapter 2, the positioning of helix 12 in the structure of the ERβ LBD bound to THC may reflect the positioning of helix 12 in the apo-ERβ LBD and perhaps other apo-SHR LBDs. When helix 12 is in this position in ERβ, one of the residues of helix 12 (Val 487) is inserted into the entrance of the ligand binding pocket formed by helices 3, 6, and 11 preventing the entry of ligands into the binding pocket. This suggests that a partial unfolding of the LBD which resulted in the removal of helix 12 from this position should permit the entry of ligands into the binding pocket.

These observations suggest the following simple model. Initially the apo-receptor LBD interacts with hsp90 in the open configuration (Figure 11). This is supported by biochemical evidence indicating that unliganded PR are able to associate with hsp90 in the absence of ATP (Kosano et al., 1998; Toft, 1999). Upon ATP binding, hsp90 undergoes a conformational change to the closed configuration. Due to its now restricted size, the binding cleft formed between the two monomers is only able to interact with helix 12 (Figure 11). Those regions of the LBD outside of helix 12 are stabilized through

interactions with regions of hsp90 outside of the binding cleft and p23 (Howard et al., 1990). The removal of helix 12 from the entrance to the binding pocket makes the LBD competent to bind hormone (Figure 11). These proposed steps are supported by experiments indicating that the addition of both ATP and p23 are necessary for the PR and GR LBDs to acquire the ability to bind steroid (Dittmar et al., 1997; Kosano et al., 1998; Toft, 1999)]. The binding of steroid and the hydrolysis of ATP would allow the release of the liganded receptor (Figure 11). Further biochemical and genetic studies will ultimately be needed to test the validity of this model.

Experimental Procedures

Cloning and Mutagenesis

Escherichia coli HtpG was expressed using a T7 RNA polymerase-based expression system. PCR (template = pBJ2 (Bardwell and Craig, 1987); primers = 5′-TTATTGAGGTAGACCCATATGAAAGGACAAGAA-3′ and 5′-GCCTCTAGATCAGGAAACCAGCAGCTGGTTCATACGACG-3′) was used to generate a fragment containing the entire HtpG open reading frame flanked by an NdeI site (at the 5′ end) and an XbaI site (at the 3′ end). The plasmid, pT7HtpGPCR, was created by ligating this insert, digested with NdeI and XbaI, with a pET-3a derivative, digested with the same enzymes. In order to minimize the probability that PCR had introduced mutations into the open reading frame, a large fragment of the of the HtpG open reading frame in pT7HtpGPCR was replaced with the equivalent fragment from pBJ2. pBJ2 was digested with EcoRI and the resulting 5′ overhangs blunted with Klenow. The fragment of interest was then released using a KpnI digest. pT7HtpGPCR was digested with XhoI and the resulting 5′ overhangs blunted with Klenow. After digesting with KpnI, the plasmid was then ligated with the insert from pBJ2. The correctness of the sequence between the NdeI and KpnI sites of the resulting expression plasmid, pT7HtpG, was then verified by manual double-stranded sequencing.

The cysteine mutants of HtpG were generated in pT7HtpG using the Chameleon site-directed mutagenesis kit (Stratagene) and the manufacturer's

instructions (with some modifications). Relative to the recommendations of the manufacturer, the amount of plasmid was decreased four-fold and the amount of mutagenic primer was increased two-fold (Primers =

S40C: 5′-CTTATCTCTAACGCGTGCGATGCGGGGAC-3′;

S79C: 5′-GTACGCTGACCATCTGCGATAACGGCGTGG-3′;

S104C: 5′-CGCTAAATCCGGCACCAAATGCTTCCTCGAATCCC-3′;

S156C: 5′-GTCTTCTGGGAATGCGCAGGCGAAGGTGAATAC-3′;

S198C: 5′-GCGTTCCATCATCTGCAAATATTCCGACATATCGC-3′;

T220C: 5′-GAAAGACGGCGAATGCGTTATCTCCTGG-3′;

S282C: 5′-GCTGTACATCCCGTGCCAGGCGCCGTGGGATATGTG-3′;

A310C: 5′-CATCATGGACGACTGCGAACAGTTCATGCC-3′;

S335C: 5′-CTGCCGCTGAACGTTTGCCGTGAAATCCTGCAGGACAGC
ACGGTAACG-3′;

V356C: 5′-GCTGACCAAGCGTTGCCTGCAGATGCTGGAAAAAC-3′;

S421C: 5′-GTATCTCTGGAAGACTACGTATGCGCATGAAAGAAGG-3′;

S437C: 5′-CTACTACATCACCGCGGACTGCTATGCGGCAGCG-3′;

S461C: 5′-GTTCTGCTGCTTTGCGATCGTATCGATGAGTGG-3′;

S484C: 5′-GTTCCAGTCGGTGTGCAAAGTGGACGAGTCGCTTG-3′;

S500C: 5′-GACGAAGTTGATGAATGCGCGAAAGAAGCGG-3′;

S547C: 5′-GCGGACGAAATGTGCACTCAGATGG-3′;

S590C: 5′-GATGAAGCGAAGTTCTGCGAGTGGGTAGAAC-3′). Successful introduction of the mutations was initially assessed either through restriction

digests (to verify the incorporation of a new restriction site) or by automated sequencing.

Expression constructs for the N517 and 497C fragments were prepared by modifying pT7HtpG. The single SalI site of pT7HtpG was destroyed by cutting pT7HtpG with SalI, filling in the overhangs with Klenow, and religating the plasmid creating pT7HtpGΔSalI. Concurrently, a ClaI/BamHI fragment (containing the 3′ end of the HtpG gene) was ligated into pBSKS+ II digested with the same enzymes creating pBSClaBam. Kunkel mutagenesis was used to introduce two silent mutations to generate a SalI site within the ClaI/BamHI fragment. After verifying the introduction of the mutations by manual double-stranded sequencing, the ClaI/BamHI fragment of the mutagenized version of pBSClaBam was ligated into pT7HtpGΔSalI digested with BamHI and partially digested with ClaI creating pT7HtpGSalI. By ligating the appropriate oligonucleotide cassettes into pT7HtpGSalI digested with SalI and BamHI or digested with NdeI and SalI, pT7N517 and pT7His$_6$497C were generated respectively.

Protein Expression

Full-length HtpG

Typically, 3-4 colonies of BL21(DE3), transformed with either pT7HtpG or its mutant derivatives, were used to inoculate culture tubes of NZ broth (10 g NZ-amine, 5 g yeast extract, and 5 g NaCl per liter) supplemented with 100 μg/mL carbenicillin. These starter cultures were shaken at least 16 hours at

250-300 rpm in an air shaker at 37 °C. Portions of these cultures were diluted to an OD_{600} of ~0.05 in NZ broth supplemented with 100 μg/mL carbenicillin. The expression cultures were shaken at 250-300 rpm at 37°C until the OD_{600} reached 0.5-1.0 (typically 2-3 hours). IPTG was added to a final concentration of 100-200 μM to induce expression. Higher concentrations of IPTG generally resulted in reduced levels of the expressed protein. After an additional 3-4 hours of shaking at 250-300 rpm at 37°C, the cells were harvested by centrifugation, frozen in liquid nitrogen, and stored at -80°C.

Selenomethionyl HtpG was expressed in B834(DE3) (auxotrophic for methionine) transformed with pT7HtpG. Colonies of B834(DE3)pT7HtpG (selected on NZ agar plates supplemented with 100 μg/mL carbenicillin) were used to inoculate culture tubes of a variant of LeMaster medium (was replaced with 5 mg/L thiamine) supplemented with 50 μg/mL seleno-DL-methionine and 100 μg/mL carbenicillin. These starter cultures were shaken at least 16 hours at 250-300 rpm in an air shaker at 37 °C. Portions of these cultures were diluted to an OD_{600} of ~0.01 in the medium described above. The expression cultures were shaken at 250-300 rpm at 37°C until the OD_{600} reached 0.6 (~9 hours). IPTG was added to a final concentration of 1 mM to induce expression. After an additional 6 hours of shaking at 250-300 rpm at 37°C, the cells were harvested by centrifugation, frozen in liquid nitrogen, and stored at -80°C.

HtpG N517

Typically, 3-4 colonies of BL21(DE3), transformed with pT7N517, were used to inoculate culture tubes of LB broth supplemented with 100 μg/mL carbenicillin. These starter cultures were shaken at least 16 hours at 250 rpm in an air shaker at 37 °C. Portions of these cultures were diluted to an OD_{600} of ~0.025 in LB broth supplemented with 100 μg/mL carbenicillin. The expression cultures were shaken at 250 rpm at 37°C until the OD_{600} reached 0.6 (typically 3-4 hours). IPTG was added to a final concentration of 200 μM to induce expression. After an additional 3 hours of shaking at 250 rpm at 37°C, the cells were harvested by centrifugation. The cell pellet was resuspended in ~50 mL 1X Phosphate Buffered Saline supplemented with 500 μM PMSF and benzamidine·HCl. The cells were again collected by centrifugation, frozen in liquid nitrogen, and stored at -80°C.

HtpG $His_6$497C

Typically, 3-4 colonies of BL21(DE3)pLysS, transformed with pT7$His_6$497C, were used to inoculate culture tubes of LB broth supplemented with 100 μg/mL carbenicillin and 20 μg/mL chloramphenicol. These starter cultures were shaken at least 16 hours at 250 rpm in an air shaker at 37 °C. Portions of these cultures were diluted to an OD_{600} of ~0.005 in LB broth supplemented

with 100 μg/mL carbenicillin and 20 μg/mL chloramphenicol. The expression cultures were shaken at 250 rpm at 37°C until the OD_{600} reached 0.6-0.7 (typically 4 hours). IPTG was added to a final concentration of 300 μM to induce expression. After an additional 3 hours of shaking at 250 rpm at 37°C, the cells were harvested by centrifugation. The cell pellet was resuspended in ~50 mL 1X Phosphate Buffered Saline supplemented with 500 μM PMSF and benzamidine·HCl. The cells were again collected by centrifugation, frozen in liquid nitrogen, and stored at -80°C.

Protein Purification

Full-length HtpG

Frozen cell pellets were resuspended in ice cold lysis buffer (10 mM Tris·Cl, 1 mM EDTA·Na_2, 500 mM NaCl, 1 mM AEBSF, 2.5 mM benzamidine·HCl, 5-10 mM DTT pH 8.0; 25 mL per g of cell paste). The cells (in a magnetically stirred, ice cold suspension) were lysed by sonication using a Branson Sonifier (constant duty cycle; standard tip probe; 15 second pulses at 60-80% maximum power with 45 second breaks; 8-10 pulses total). The crude lysate was clarified by centrifugation at 45,000 rpm in a Ti45 rotor (Beckman) for 1 hour at 4°C. The total protein concentration of the extract was determined by Bradford assay and the extract was diluted with lysis buffer until the total protein concentration was ~ 2 mg/mL. An ice cold solution of protamine sulfate (concentration = 0.25 x (total protein concentration)/6) was added dropwise

with stirring to the extract. The resulting suspension was centrifuged at 15,000 rpm in an SS-34 rotor (Sorvall) for 0.5-1hour at 4°C. Solid ammonium sulfate was added slowly to the supernatant to a final concentration of 70%. The precipitated protein was harvested by centrifugation (15,000 rpm in an SS-34 rotor for 1 hour at 4°C). The protein pellet was dissolved by gentle resuspension in loading buffer (20 mM Tris·Cl, 10 mM DTT pH 8.0; 1.5 mL per mg of protein). The resulting protein solution was clarified either by filtration through a 0.45 μm filter (Millipore) or by centrifugation at 15,000 rpm in an SS-34 rotor (Sorvall) for 0.5 hour at 4°C. The sample was loaded onto a HiTrap Q column (Pharmacia) equilbrated in the loading buffer. The column was washed with 5-10 column volumes of the loading buffer followed by 5-10 column volumes of column buffer (10 mM Bis-Tris·Cl, 100 mM NaCl, 5 mM DTT pH 6.0). The column was developed with a 100 mM -> 350-400 mM NaCl gradient (20-30 column volumes) in column buffer. Wild-type HtpG elutes at ~200 mM NaCl and the mutant proteins generally eluted within +/- 50 mM of this value. Fractions containing the protein were concentrated to 20-150 mg/mL and buffer exchanged into a storage buffer (10 mM Tris·Cl, 5-10 mM DTT pH 7.5-8.0 or 10 mM HEPES·Na, 5-10 mM DTT pH 7.5) by ultrafiltration (using a combination of Centriprep 30s, Centricon 30s, and Centricon 50s). The protein (typically 90-95% pure) was aliquoted, frozen in liquid nitrogen, and stored at -80°C.

Selenomethionyl HtpG was purified essentially as described above except that all buffers were supplemented with 10 mM L-methionine after exhaustive vacuum degassing.

HtpG N517

Frozen cell pellets were resuspended in ice cold lysis buffer (20 mM Bis-Tris·Cl, 1 mM EDTA·Na_2, 1 mM AEBSF, 2.5 mM benzamidine·HCl, 5 mM DTT pH 6.8; 25 mL per g of cell paste). The cells (in a magnetically stirred, ice cold suspension) were lysed by sonication using a Branson Sonifier (constant duty cycle; standard tip probe; 30 second pulses at 60 maximum power with 60 second breaks; 6 pulses total). The crude lysate was clarified by centrifugation at 45,000 rpm in a Ti45 rotor (Beckman) for 1 hour at 4°C. The supernatant was then filtered through a 0.45 μm filter (Millipore). The sample was loaded onto a column of Fast Flow Q Sepharose(Pharmacia) equilbrated in the lysis buffer. The column was washed with 3 column volumes of the lysis buffer and then developed with a 0 mM -> 500 mM NaCl gradient (10 column volumes). HtpG N517 elutes at ~200 mM NaCl. Fractions containing the protein were dialyzed against 50 mM Tris·Cl, 50 mM KCl, 1 mM $CaCl_2$, 2 mM DTT pH 7.8 and loaded onto a column of hydroxyapatite Spectragel (Spectrum) equilbrated in the dialysis buffer without calcium. The column was washed with 5-10 column volumes of the dialysis buffer without calcium and developed with a 0 mM -> 250 mM potassium phosphate gradient (20 column volumes) in the dialysis buffer without calcium. HtpG N517 elutes at

~100 mM potassium phosphate. Fractions containing the protein were pooled, dialyzed into 5 mM Tris·Cl, 1 mM DTT pH 8.0, concentrated to 33 mg/mL by ultrafiltration using Centricon 30s. The protein (~90-95% pure) was aliquoted, frozen in liquid nitrogen, and stored at -80°C.

HtpG $His_6$497C

Frozen cell pellets were resuspended in ice cold lysis buffer (20 mM Tris·Cl, 500 mM NaCl, 1 mM Imidazole pH 8.0; 25 mL per g of cell paste). The cells (in a magnetically stirred, ice cold suspension) were lysed by sonication using a Branson Sonifier (constant duty cycle; standard tip probe; 30 second pulses at 60% maximum power with 60 second breaks; 4-5 pulses total). The crude lysate was clarified by centrifugation at 15,000 rpm in an SS-34 rotor (Sorvall) for 20 minutes at 4°C. The supernatant was clarified either by filtration through a 0.2 μm filter (Millipore) and then loaded onto a column of Chelating Fast Flow Sepharose (Pharmacia) (~2.5 mL per g of cell paste) charged with Co^{++} and equilbrated in lysis buffer. The column was washed with 10 column volumes of the lysis buffer and then developed with a 1 mM -> 200 mM Imidazole gradient (20 column volumes). $His_6$497C elutes as a broad peak starting at ~75 mM Imidazole. Fractions containing the protein were dialyzed against 10 mM Tris·Cl, 1 mM EDTA·Na_2 pH 8.0 and loaded onto a HiTrap Q column (Pharmacia) equilbrated in the loading buffer. The column was washed with 10 column volumes of the dialysis buffer and developed with a 0 mM -> 250 mM NaCl gradient (20 column volumes) in

the dialysis buffer. Fractions containing the protein were concentrated to ~20 mg/mL and buffer exchanged into a storage buffer (10 mM Tris·Cl, 100 mM NaCl, 1 mM EDTA·Na_2 pH 8.0) by ultrafiltration using Centricon 10s. The protein (~90-95% pure) was aliquoted, frozen in liquid nitrogen, and stored at -80°C.

Lucifer Yellow Labeling

Frozen cell pellets from 5 mL cultures were resuspended in 450 µL ice cold phosphate lysis buffer (50 mM sodium phosphate, 1 mM EDTA·Na_2, 200 mM NaCl, 1 mM AEBSF, 2.5 mM benzamidine·HCl, 1 mM DTT pH 7.2). The cells were lysed by sonication using a Branson Sonifier (constant duty cycle; microtip probe; 5-10 second pulses at 20% maximum power with 5-10 second breaks; 3 pulses total). The crude lysate was centrifuged at 14,000 rpm in a microfuge (Eppendorf) for 1 hour at 4°C and the total protein concentration of the supernatant determined by a Bradford assay.

The protein in a 200 µL sample of the clarified extract was denatured by the addition of 200 mg of solid urea and 50 µL of water (final volume ~400 µL; final urea concentration ~8 M). After 1 hour at room temperature, 40 µg of the urea denatured protein was diluted with urea buffer (8 M urea, 25 mM sodium phosphate, 0.5 mM EDTA·Na_2, 100 mM NaCl pH 7.2) to a final volume of 30 µL. The protein sample was split into two 15 µL aliquots. Five µL of urea buffer was added to one aliquot and 5 µL of urea buffer containing

16 mM Lucifer Yellow iodoacetamide (Molecular Probes) was added to the other. After 1 hour at room temperature in the dark, the labeling reactions were quenched with DTT (final concentration = 50 mM). A portion of each reaction (3.5 μg total protein) was subjected to SDS-PAGE on 4-12% Tris-glycine gels. The proteins were visualized first by UV transillumination and then by staining with Coomassie dye.

Similar reactions were performed on samples of protein in the native state. An aliquot (40 μg total protein) of the clarified extract was diluted with phosphate lysis buffer (without protease inhibitors) to a final volume of 30 μL. The sample was split into two 15 μL aliquots. Five μL of phosphate lysis buffer containing 6.4 mM Lucifer Yellow iodoacetamide was added to each. One of the labeling reactions was quenched with DTT (final concentration = 50 mM) after 1 hour at room temperature in the dark and the other was quenched after 20 hours at room temperature in the dark. A portion of each reaction (3.5 μg total protein) was subjected to SDS-PAGE on 4-12% Tris-glycine gels. The proteins were visualized first by UV transillumination and then by staining with Coomassie dye.

Size Exclusion Chromatography and Crosslinking

Protein samples (50 μL at 0.05 to 1 mg/mL) were analyzed by size exclusion chromatography on a Superose 12 column (Pharmacia) equilibrated in 0.5 M

NH_4HCO_3 pH 8.0. Molecular weights were estimated using commercially available standards (Biorad) for calibration.

Prior to crosslinking, purified full-length HtpG, N517, and $His_6$497C were diluted to 100 μg/mL in 40 mM Sodium Phosphate, 400 mM NaCl pH 6.8. Samples (5 μL) of these protein solutions were combined with an equal volume of water or 0.02%, 0.1%, or 0.2% glutaraldehyde in water. The reactions were incubated at room temperature for fifteen minutes and then quenched by the addition of 0.5 μL of 1 M lysine in water. The reactions were subjected to SDS-PAGE on 12% Tris-glycine gels and the proteins visualized by staining with Coomassie blue.

Derivatization of Cysteine Mutants

A frozen aliquot (40-100 μL) containing 2-4 mg of protein was rapidly thawed and diluted with 2 mL of 50 mM HEPES·Na pH 7.5. The protein sample was concentrated in a Centricon 30 (previously rinsed with water) to a volume of ~200 μL. The protein sample was again diluted with 2 mL of 50 mM HEPES·Na pH 7.5 and concentrated to a volume of ~200 μL. Methyl mercury nitrate was prepared as previously described (Nagai et al., 1991). Two mL of 50 mM HEPES·Na with 1 mM methyl mercury nitrate pH 7.5 was added and the reaction was allowed to incubate in the dark at 4°C for 1 hour. The reaction was repeatedly buffer exchanged with 10 mM HEPES·Na pH 7.5 to reduce the final concentration of the mercury compound to less than 1 μM. The reacted

protein was concentrated to ~15-40 mg/mL, aliquoted, frozen in liquid nitrogen, and stored at -80 °C.

Crystallization

Crystals were obtained by hanging drop vapor diffusion at 19-23 °C. Samples (2-3 μL) of the precipitant solution (5-10% (w/v) PEG 3000, 30-60 mM $MgCl_2$, 0.25% β-octylglucoside, 50 mM Tris·Cl pH 7.5-9.0) were first pipetted onto siliconized cover slides. Samples (2-3 μL) of 12-13 mg/mL solutions of either wild-type HtpG or the cysteine mutants were then gently layered on top of these droplets of precipitant. The solutions of underivatized cysteine mutants contained 5 mM DTT. Finally, the cover slides were rapidly inverted and suspended over wells containing 0.5-1 mL of well buffer (5-10% (w/v) PEG 3000, 30-60 mM $MgCl_2$, 50 mM Tris·Cl pH 7.5-9.0). Under these conditions, hexagonal bipyramidal crystals generally formed within 5 minutes and continued to grow for 3-4 days to a maximum size of 0.5 mm x 0.5 mm x 1 mm. Any deviations from this protocol, such as reversing the order of addition of the precipitant and protein to the cover slide, led to either the formation of showers of microcrystals or the formation of amorphous precipitates. These crystals lie in the spacegroup $P6_522$ with a = b = 105.06 Å and c = 531.62 Å for crystals grown at pH 8 and a = b = 104.66 Å and c = 525.73 Å for crystals grown at pH 9. By Ficoll gradient density measurements (Westbrook, 1985), two molecules of HtpG form the asymmetric unit.

Flash Cooling and Data Collection

In preparation for cryocooling, the crystals were first transferred to a protein-free stabilizing solution (10% (w/v) PEG 3000, 50 mM $MgCl_2$, 0.50% β-octylglucoside, 50 mM Tris·Cl pH 8.0-9.0). The crystals were then sequentially transferred at 1 minute intervals through a series of solutions consisting of the stabilizing solution supplemented with increasing concentrations of ethylene glycol (5%, 10%, 15%, and 20%). The crystals were either frozen in a stream of nitrogen gas or in a bath of liquid nitrogen.

Data were measured from these crystals at –170°C either at SSRL beamline 1-5 using Fuji BAS-2000 image plates at a wavelength of 1.54 Å or at CHESS station F-1 using a 2 X 2 CCD camera or Fuji BAS-2000 image plates at a wavelength of 0.92 Å. Data were reduced with the DENZO/Scalepack package (Otwinowski and Minor, 1997) using the default –3σ cutoff.

Programs

Heavy atom sites were determined initially determined by manual inspection of difference Patterson maps and later difference Fourier maps calculated in XtalView (McRee, 1999). Heavy atom parameters were refined and MIR phases calculated in MLPHARE (CCP4, 1994). All masks were generated and modified using MAMA (Kleywegt and Jones, 1994). Multicrystal form averaging was performed with MAVE (Jones, 1992) and NCS averaging was

performed using SOLOMON (CCP4, 1994). All model building was using MOLOC (Muller et al., 1988).

Illustrations

Figures 5, 6, 7, 8, and 10 were generated using BOBSCRIPT (Esnouf, 1997) and rendered using Raster3D (Merritt and Anderson, 1994).

References

Allan, G. F., Leng, X., Tsai, S. Y., Weigel, N. L., Edwards, D. P., Tsai, M. J., and O'Malley, B. W. (1992). Hormone and antihormone induce distinct conformational changes which are central to steroid receptor activation. J. Biol. Chem. *267*, 19513-20.

Ban, C., Junop, M., and Yang, W. (1999). Transformation of MutL by ATP binding and hydrolysis: a switch in DNA mismatch repair. Cell *97*, 85-97.

Ban, C., and Yang, W. (1998). Crystal structure and ATPase activity of MutL: implications for DNA repair and mutagenesis. Cell *95*, 541-52.

Bardwell, J. C., and Craig, E. A. (1988). Ancient heat shock gene is dispensable. J. Bacteriol. *170*, 2977-83.

Bardwell, J. C., and Craig, E. A. (1987). Eukaryotic Mr 83,000 heat shock protein has a homologue in Escherichia coli. Proc. Natl. Acad. Sci. U S A *84*, 5177-81.

Beekman, J. M., Allan, G. F., Tsai, S. Y., Tsai, M. J., and O'Malley, B. W. (1993). Transcriptional activation by the estrogen receptor requires a conformational change in the ligand binding domain. Mol. Endocrinol. *7*, 1266-74.

Blundell, T. L., and Johnson, L. N. (1976). Protein Crystallography (London: Academic Press).

Bohen, S. P., and Yamamoto, K. R. (1994). Modulation of steroid receptor signal transduction by heat shock proteins. In The Biology of Heat Shock Proteins and Molecular Chaperones, R. I. Morimoto, A. Tissiares and C. Georgopoulos, eds. (Cold Spring Harbor: Cold Spring Harbor Laboratory).

Borkovich, K. A., Farrelly, F. W., Finkelstein, D. B., Taulien, J., and Lindquist, S. (1989). Hsp82 is an essential protein that is required in higher concentrations for growth of cells at higher temperatures. Mol. Cell Biol. *9*, 3919-30.

Brzozowski, A., Pike, A., Dauter, Z., Hubbard, R., Bonn, T., Engstrom, O., Ohman, L., Greene, G., Gustafsson, J., and Carlquist, M. (1997). Molecular basis of agonism and antagonism in the oestrogen receptor. Nature *389*, 753-758.

Buchner, J. (1999). Hsp90 & co. - a holding for folding. Trends Biochem. Sci. *24*, 136-41.

Bukau, B. (1999). Molecular Chaperones and Folding Catalysts: Regulation, Cellular Function, and Mechanisms (Amsterdam: Harwood Academic Publishers).

Bukau, B., and Horwich, A. L. (1998). The hsp70 and hsp60 chaperone machines. Cell *92*, 351-66.

CCP4 (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D *50*, 760-763.

Cohen, G. N., and Saint-Girons, I. (1987). Biosynthesis of threonine, lysine, and methionine. In Escherichia coli and Salmonella typhimurium, F. C. Neidhardt, J. L. Ingraham, B. Low, B. Magasanik, M. Schaechter and H. E. Umbarger, eds. (Washington, D. C.: American Society for Microbiology), pp. 429-444.

Cutforth, T., and Rubin, G. M. (1994). Mutations in hsp83 and cdc37 impair signaling by the sevenless receptor tyrosine kinase in Drosophila. Cell *77*, 1027-36.

Dittmar, K. D., Demady, D. R., Stancato, L. F., Krishna, P., and Pratt, W. B. (1997). Folding of the glucocorticoid receptor by the heat shock protein (hsp) 90-based chaperone machinery. The role of p23 is to stabilize receptor.hsp90 heterocomplexes formed by hsp90.p60.hsp70. J. Biol. Chem. 272, 21213-20.

Esnouf, R. M. (1997). An extensively modified version of MolScript that includes greatly enhanced coloring capabilities. J. Mol. Graph. Model. *15*, 132-4, 112-3.

Freeman, B. C., Toft, D. O., and Morimoto, R. I. (1996). Molecular chaperone machines: chaperone activities of the cyclophilin Cyp-40 and the steroid aporeceptor-associated protein p23. Science *274*, 1718-20.

Gupta, R. S. (1995). Phylogenetic analysis of the 90 kD heat shock family of protein sequences and an examination of the relationship among animals, plants, and fungi species. Mol. Biol. Evol. *12*, 1063-73.

Hartl, F. U. (1996). Molecular chaperones in cellular protein folding. Nature *381*, 571-9.

Hatfull, G. F., Sanderson, M. R., Freemont, P. S., Raccuia, P. R., Grindley, N. D., and Steitz, T. A. (1989). Preparation of heavy-atom derivatives using site-directed mutagenesis. Introduction of cysteine residues into gamma delta resolvase. J. Mol. Biol. *208*, 661-7.

Herendeen, S. L., VanBogelen, R. A., and Neidhardt, F. C. (1979). Levels of major proteins of Escherichia coli during growth at different temperatures. J. Bacteriol. *139*, 185-94.

Holley, S. J., and Yamamoto, K. R. (1995). A role for hsp90 in retinoid receptor signal transduction. Mol. Biol. Cell *6*, 1833-42.

Howard, K. J., Holley, S. J., Yamamoto, K. R., and Distelhorst, C. W. (1990). Mapping the hsp90 binding region of the glucocorticoid receptor. J. Biol. Chem. *265*, 11928-35.

Jones, T. A. (1992). A, yaap, asap, @#*? A set of averaging programs. In Molecular Replacement, E. J. Dodson, S. Gover and W. Wolf, eds. (Warrington: SERC Daresbury Laboratory), pp. 91-105.

Kallio, P. J., Janne, O. A., and Palvimo, J. J. (1994). Agonists, but not antagonists, alter the conformation of the hormone- binding domain of androgen receptor. Endocrinology *134*, 998-1001.

Kleywegt, G. J., and Jones, T. A. (1994). Halloween...masks and bones. In From First Map to Final Model, S. Bailey, R. Hubbard and D. Waller, eds. (Warrington, England: SERC Daresbury Laboratory).

Koch, G., Smith, M., Macer, D., Webster, P., and Mortara, R. (1986). Endoplasmic reticulum contains a common, abundant calcium-binding glycoprotein, endoplasmin. J. Cell Sci. *86*, 217-32.

Kosano, H., Stensgard, B., Charlesworth, M. C., McMahon, N., and Toft, D. (1998). The assembly of progesterone receptor-hsp90 complexes using purified proteins. J. Biol. Chem. *273*, 32973-9.

Krone, P. H., and Sass, J. B. (1994). Hsp90 alpha and hsp90 beta genes are present in the zebrafish and are differentially regulated in developing embryos. Biochem. Biophys. Res. Commun. *204*, 746-52.

LeMaster, D. M., and Richards, F. M. (1985). 1H-15N heteronuclear NMR studies of Escherichia coli thioredoxin in samples isotopically labeled by residue type. Biochemistry *24*, 7263-8.

Maruya, M., Sameshima, M., Nemoto, T., and Yahara, I. (1999). Monomer arrangement in hsp90 dimer as determined by decoration with N and C-terminal region specific antibodies. J. Mol. Biol. *285*, 903-7.

McRee, D. E. (1999). XtalView/Xfit--A versatile program for manipulating atomic coordinates and electron density. J. Struct. Biol. *125*, 156-65.

Melnick, J., Dul, J. L., and Argon, Y. (1994). Sequential interaction of the chaperones BiP and grp94 with immunoglobulin chains in the endoplasmic reticulum. Nature *370*, 373-5.

Meng, X., Devin, J., Sullivan, W. P., Toft, D., Baulieu, E. E., and Catelli, M. G. (1996). Mutational analysis of hsp90 alpha dimerization and subcellular localization: dimer disruption does not impede "in vivo" interaction with estrogen receptor. J. Cell Sci. *109*, 1677-87.

Merritt, E. A., and Anderson, W. F. (1994). Raster3D Version 2.0: A program for photorealistic molecular structures. Acta Crystallogr. D *50*, 219-220.

Minami, Y., Kimura, Y., Kawasaki, H., Suzuki, K., and Yahara, I. (1994). The carboxy-terminal region of mammalian hsp90 is required for its dimerization and function in vivo. Mol. Cell Biol. *14*, 1459-64.

Modarress, K. J., Opoku, J., Xu, M., Sarlis, N. J., and Simons, S. S., Jr. (1997). Steroid-induced conformational changes at ends of the hormone-binding domain in the rat glucocorticoid receptor are independent of agonist versus antagonist activity. J. Biol. Chem. *272*, 23986-94.

Muller, K., Amman, H. J., Doran, D. M., Gerber, P. R., Gubernator, K., and Schrepfer, G. (1988). MOLOC: A molecular modeling program. Bull. Soc. Chim. Belg. *97*, 655-667.

Nagai, K., Evans, P. R., Li, J., and Oubridge, C. (1991). Phase determination using mercury derivatives of engineered cysteine mutants. In Proc. CCP4 Study Weekend, P. R. Evans and A. Leslie, eds. (UK: SERC Daresbury Laboratory), pp. 141-149.

Obermann, W. M., Sondermann, H., Russo, A. A., Pavletich, N. P., and Hartl, F. U. (1998). In vivo function of hsp90 is dependent on ATP binding and ATP hydrolysis. J. Cell Biol. *143*, 901-10.

Otwinowski, Z., and Minor, W. (1997). Processing of x-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307-326.

Panaretou, B., Prodromou, C., Roe, S. M., O'Brien, R., Ladbury, J. E., Piper, P. W., and Pearl, L. H. (1998). ATP binding and hydrolysis are essential to the function of the hsp90 molecular chaperone in vivo. EMBO J. *17*, 4829-36.

Picard, D., Khursheed, B., Garabedian, M. J., Fortin, M. G., Lindquist, S., and Yamamoto, K. R. (1990). Reduced levels of hsp90 compromise steroid receptor action in vivo. Nature *348*, 166-8.

Pike, A. C., Brzozowski, A. M., Hubbard, R. E., Bonn, T., Thorsell, A. G., Engstrom, O., Ljunggren, J., Gustafsson, J. A., and Carlquist, M. (1999).

Structure of the ligand-binding domain of oestrogen receptor beta in the presence of a partial agonist and a full antagonist. EMBO J. *18*, 4608-18.

Prodromou, C., Roe, S. M., O'Brien, R., Ladbury, J. E., Piper, P. W., and Pearl, L. H. (1997). Identification and structural characterization of the ATP/ADP-binding site in the hsp90 molecular chaperone. Cell *90*, 65-75.

Prodromou, C., Roe, S. M., Piper, P. W., and Pearl, L. H. (1997). A molecular clamp in the crystal structure of the N-terminal domain of the yeast hsp90 chaperone. Nat. Struct. Biol. *4*, 477-82.

Rayment, I. (1985). Treatment and manipulation of crystals. Methods Enzymol. *114*, 136-141.

Rodgers, D. W. (1997). Practical cryocrystallography. Methods Enzymol. *276*, 183-203.

Rutherford, S. L., and Lindquist, S. (1998). Hsp90 as a capacitor for morphological evolution. Nature *396*, 336-42.

Scheibel, T., and Buchner, J. (1997). The hsp90 family-an overview. In Guidebook to Molecular Chaperones and Protein-Catalysts, M. J. Gething, ed. (Oxford: Oxford University Press), pp. 147-151.

Schulz, A., Schwab, S., Homuth, G., Versteeg, S., and Schumann, W. (1997). The HtpG gene of Bacillus subtilis belongs to class III heat shock genes and is under negative control. J Bacteriol. *179*, 3103-9.

Seielstad, D., Carlson, K., Katzenellenbogen, J., Kushner, P., and Greene, G. (1995). Molecular characterization by mass spectrometry of the human estrogen receptor ligand-binding domain expressed in Escherichia coli. Mol. Endocrinol. *9*, 647-658.

Simons, S. S., Jr., Sistare, F. D., and Chakraborti, P. K. (1989). Steroid binding activity is retained in a 16-kDa fragment of the steroid binding domain of rat glucocorticoid receptors. J. Biol. Chem. *264*, 14493-7.

Spence, J., and Georgopoulos, C. (1989). Purification and properties of the Escherichia coli heat shock protein, HtpG. J. Biol. Chem. *264*, 4398-403.

Stebbins, C. E., Russo, A. A., Schneider, C., Rosen, N., Hartl, F. U., and Pavletich, N. P. (1997). Crystal structure of an hsp90-geldanamycin complex: targeting of a protein chaperone by an antitumor agent. Cell *89*, 239-50.

Sun, D. P., Alber, T., Bell, J. A., Weaver, L. H., and Matthews, B. W. (1987). Use of site-directed mutagenesis to obtain isomorphous heavy-atom derivatives

for protein crystallography: cysteine-containing mutants of phage T4 lysozyme. Protein Eng. *1*, 115-23.

Tanaka, N., and Nakamoto, H. (1999). HtpG is essential for the thermal stress management in cyanobacteria. FEBS Lett. *458*, 117-23.

Tanenbaum, D. M., Wang, Y., Williams, S. P., and Sigler, P. B. (1998). Crystallographic comparison of the estrogen and progesterone receptor's ligand binding domains. Proc. Natl. Acad. Sci. U S A *95*, 5998-6003.

Toft, D. O. (1999). Control of hormone receptor function by molecular chaperones and folding catalysts. In Molecular Chaperones and Folding Catalysts: Regulation, Cellular Function, and Mechanisms, B. Bukau, ed. (Amsterdam: Harwood Academic Publishers), pp. 313-327.

Wearsch, P. A., and Nicchitta, C. V. (1996). Endoplasmic reticulum chaperone grp94 subunit assembly is regulated through a defined oligomerization domain. Biochemistry *35*, 16760-9.

Westbrook, E. M. (1985). Crystal density measurements using aqueous ficoll solutions. Methods Enzymol. *114*, 187-196.

Whitesell, L., Mimnaugh, E. G., De Costa, B., Myers, C. E., and Neckers, L. M. (1994). Inhibition of heat shock protein hsp90-pp60v-src heteroprotein complex formation by benzoquinone ansamycins: essential role for stress proteins in oncogenic transformation. Proc. Natl. Acad. Sci. U S A *91*, 8324-8.

Wigley, D. B., Davies, G. J., Dodson, E. J., Maxwell, A., and Dodson, G. (1991). Crystal structure of an N-terminal fragment of the DNA gyrase B protein. Nature *351*, 624-9.

Williams, S., and Sigler, P. (1998). Atomic structure of progesterone complexed with its receptor. Nature *393*, 392-395.

Xiao, H., and Lis, J. T. (1989). Heat shock and developmental regulation of the Drosophila melanogaster hsp83 gene. Mol. Cell Biol. *9*, 1746-53.

Figure Legends

Figure 1 Crystals of HtpG. Form I crystals (up to 1 mm x 0.5 mm x 0.01 mm in size) were obtained using 1.5-2 M $(NH_4)_2SO_4$ 100 mM KOAc pH 4.5-5.5 as the precipitant. Form II crystals (up to 200 μm x 50 μm x 50 μm in size) were obtained in the presence of 1 μM geldanamycin using 2 M $(NH_4)_2SO_4$ 100 mM KOAc pH 5.0 as the precipitant. Form III crystals (up to 0.5 mm x 0.5 mm x 1.0 mm in size) were obtained using 5-10% (w/v) PEG 3000, 30-60 mM $MgCl_2$, 0.25% β-octylglucoside, 50 mM Tris·Cl pH 7.5-9.0 as the precipitant.

Figure 2 Precession Photograph. A 16 hour, 7° screened precession photograph of the hh0l reciprocal lattice section of a Form III crystal (crystal-to-detector distance = 100 mm, λ = 1.54 Å). The six-fold screw axis is readily apparent.

Figure 3 Oscillation Photograph. A 30 minute, 1° oscillation photograph of a Form III crystal collected at the SSRL Beamline 1-5 (crystal-to-detector distance = 417 mm, λ = 1.54 Å).

Figure 4 Biochemical Characterization of HtpG Cysteine Mutants.

A. Expression Levels of HtpG Cysteine Mutants. The proteins in clarified extracts of cells expressing either the wild-type or the indicated mutant HtpGs

were denatured in 8 M urea, separated by SDS-PAGE on a 10% acrylamide gel, and then visualized by staining with Coomassie blue. The wild-type protein and many of the mutants are highly expressed (accounting for ~80% of the protein in the extracts).

B. Lucifer Yellow Labeling of Denatured Proteins. The proteins in clarified extracts of cells expressing either the wild-type or the indicated mutant HtpGs were denatured in 8 M urea, labeled with lucifer yellow iodoacetamide, separated by SDS-PAGE on a 10% acrylamide gel, and then visualized by illumination with short wavelength UV light. The expressed mutant proteins are clearly labeled by with the dye, whereas the wild-type protein (which has no cysteines) is not.

C. and D. Lucifer Yellow Labeling of Native Proteins. The proteins in clarified extracts of cells expressing the wild-type or the indicated mutant HtpGs were labeled with lucifer yellow iodoacetamide for 1 hour (C.) or 20 hours (D.), separated by SDS-PAGE on a 10% acrylamide gel, and then visualized by illumination with short wavelength UV light. The labeling of the various mutant proteins occurs at different rates.

Figure 5 Experimental Electron Density of the HtpG Monomer. A stereo view of an electron density map calculated to 3.7 Å resolution using anisotropically-sharpened amplitudes and phases derived from NCS averaging. A Cα trace of the regions of the structure which have been modeled is shown in yellow. Regions of the electron density corresponding

to the N-terminal, Middle, and C-terminal domains of HtpG are boxed and labeled N, M, and C respectively.

Figure 6 N-terminal Domains of HtpG and Hsp82. Two equivalent orthogonal views of the N-terminal domains of HtpG and the yeast hsp82. Both domains are depicted as Cα worm models. For HtpG, residues 4 to 28 are colored green and residues 95 to 127 are colored magenta. For hsp82, residues 4 to 27 are colored green and residues 94 to 124 are colored magenta. A spacefilling model of the ADP bound to hsp82 is shown in yellow. Note the radically different positions of the regions colored in green in magenta which form the lid to the nucleotide binding site.

Figure 7 Middle Domains of HtpG and MutL. Two equivalent orthogonal views of a portion of the Middle domains of HtpG and MutL. Both fragments are depicted as Cα worm models. Although both adopt similar α/β/α folds, the helix connecting the four and fifth strands (colored in red) of this fragment of MutL is simply a loop (also colored in red) in the equivalent fragment of HtpG.

Figure 8 Quaternary Structures of Apo-HtpG and the ADPPNP-MutL Complex. Two equivalent orthogonal views of the HtpG dimer and the ADPPNP-MutL dimer. Both molecules are depicted as Cα worm models with one monomer colored yellow and the other colored green. The two

monomers of HtpG interact exclusively through their C-termini, which form a four-helix bundle. In contrast, the two monomers of MutL interact extensively through their N-terminal and Middle domains. As a result, the large cleft formed by the HtpG dimer is absent in the MutL dimer.

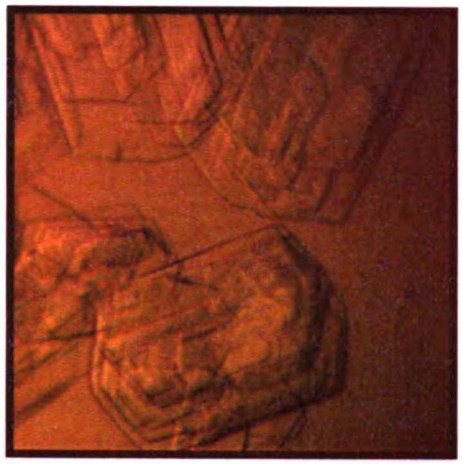
Figure 9 Crosslinking of Full-length HtpG, N517, and 497C. Samples of purified full-length HtpG, N517, and 497C were treated with the indicated concentrations of glutaraldehyde, electrophoresed on 12% acrylamide gels, and visualized by staining with Coomassie blue. Dimeric crosslinking products are evident in the full-length and 497C reactions but not in the N517 reactions.

Figure 10 Hypothetical Interaction Between HtpG and the ERβ LBD. Cα worm models of HtpG and the ERβ LBD. One of the monomers of HtpG is colored yellow and the other is colored green. The ERβ LBD is colored primarily in gray with helix 12 in magenta. The cleft formed between the two monomers of the HtpG dimer is large enough to fit a protein the size of the ERβ LBD.

Figure 11 Model for the Assembly of the Hsp90/SHR LBD/P23 Complex. In the absence of nucleotide, hsp90 (gold) interacts loosely with an apo-SHR LBD (gray) as depicted in Figure 3-10. Helix 12 of the SHR LBD (magenta) is in a

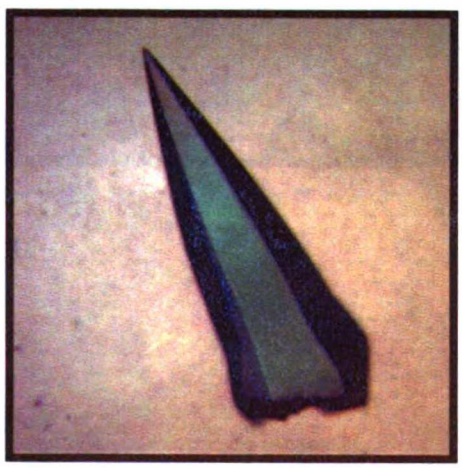
conformation resembling that seen in the THC and GEN complexes. Once ATP binds, hsp90 undergoes a conformational change that allows the N-terminal and Middle domains of the two monomers to interact in the manner seen in the ADPPNP-MutL dimer structure. Helix 12 is bound in the resulting channel formed between the monomers and the LBD is stabilized in the complex through interactions with p23 (orange). The LBD is now competent to bind hormone. The hydrolysis of ATP allows the LBD to be released.

Figure 1



Form I

Form II



Form III

Figure 2

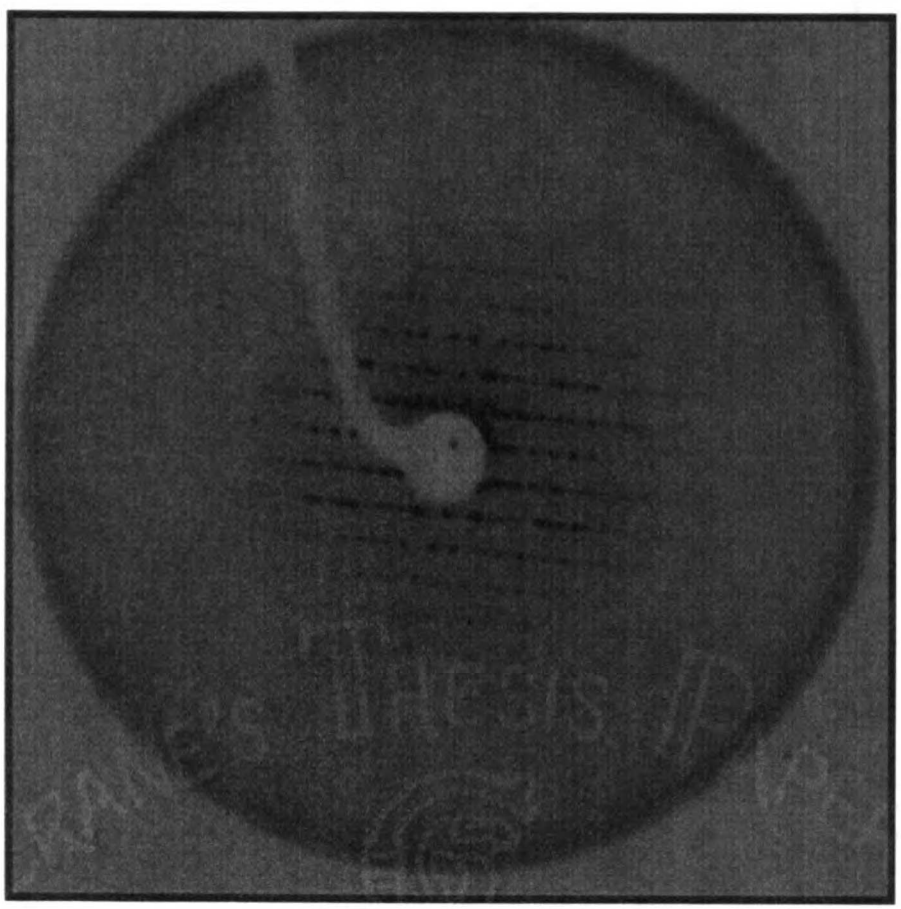


Figure 3

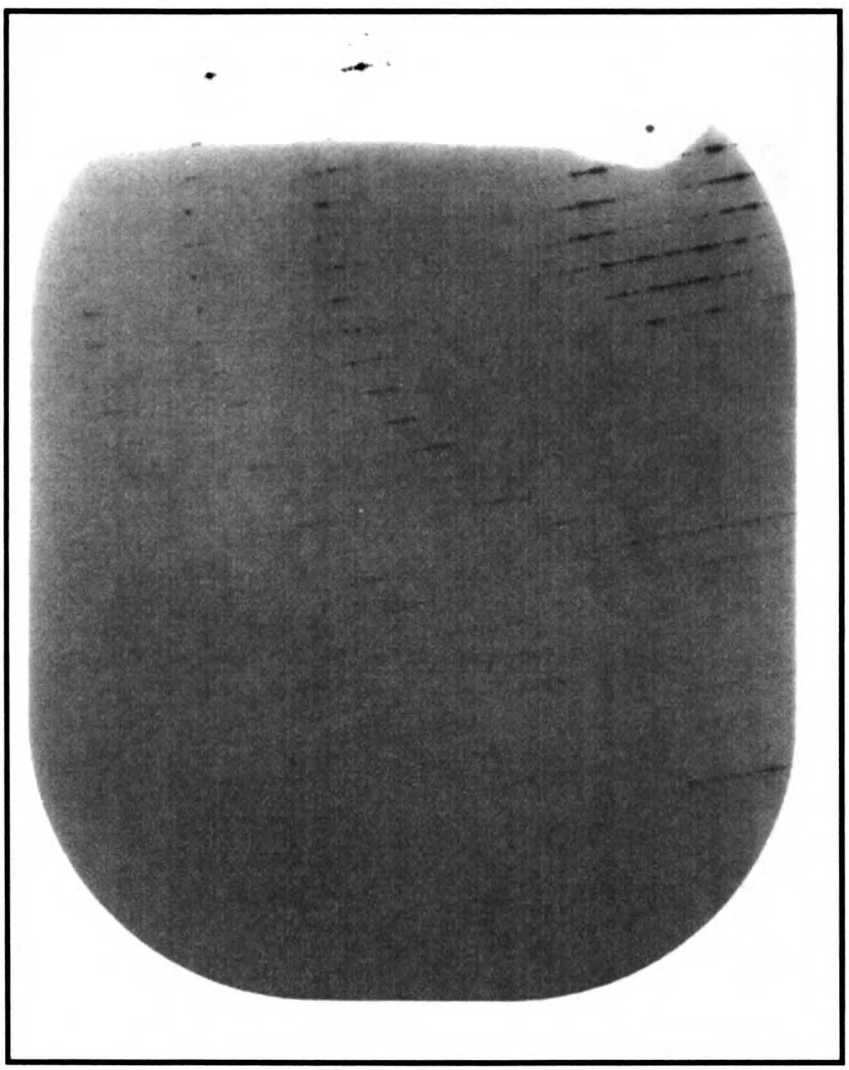


Figure 4

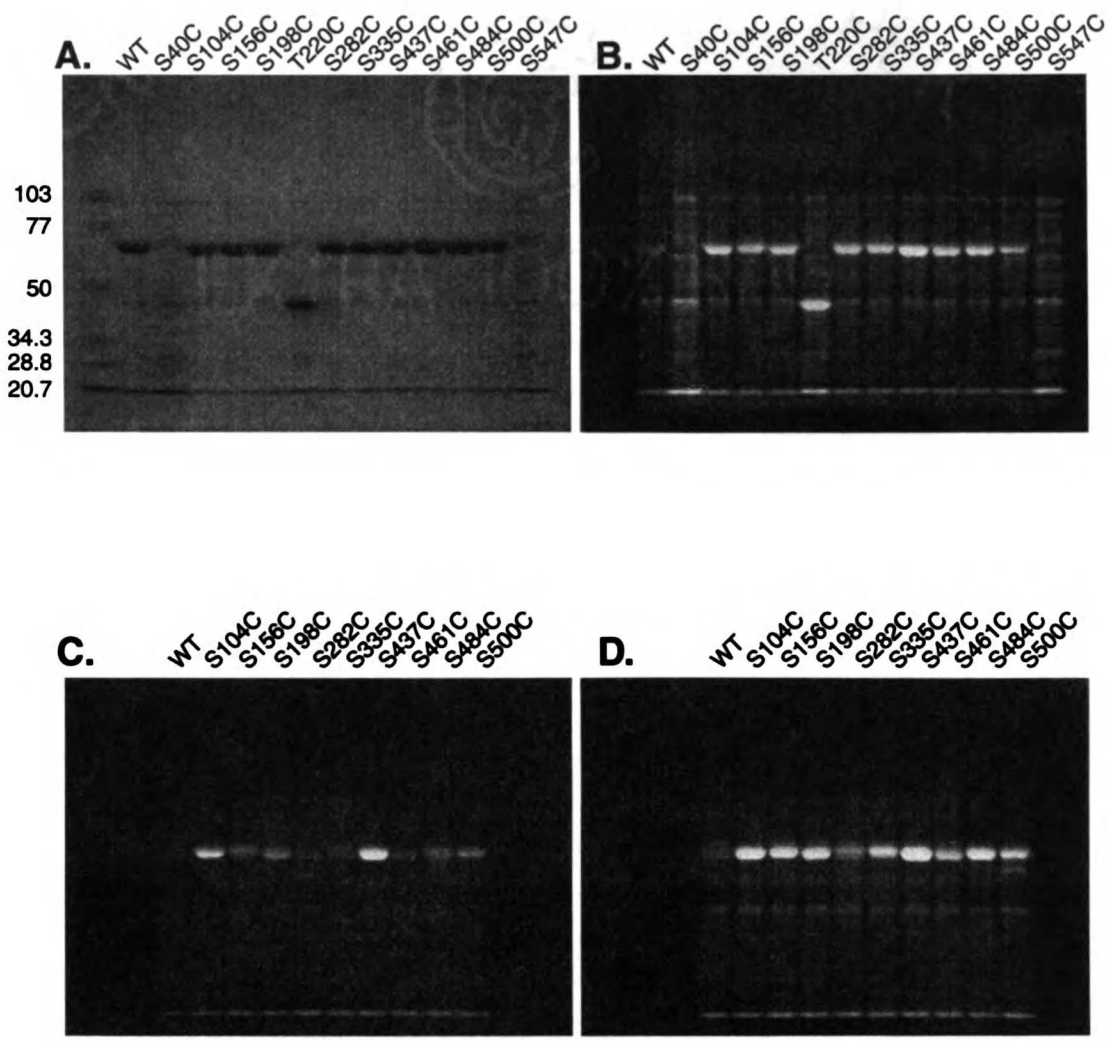

A.
WT
S40C
S104C
S156C
S198C
T220C
S282C
S335C
S437C
S461C
S484C
S500C
S547C
103
77
50
34.3
28.8
20.7
B.
WT
S40C
S104C
S156C
S198C
T220C
S282C
S335C
S437C
S461C
S484C
S500C
S547C
C.
WT
S104C
S156C
S198C
S282C
S335C
S437C
S461C
S484C
S500C
D.
WT
S104C
S156C
S198C
S282C
S335C
S437C
S461C
S484C
S500C

Figure 5

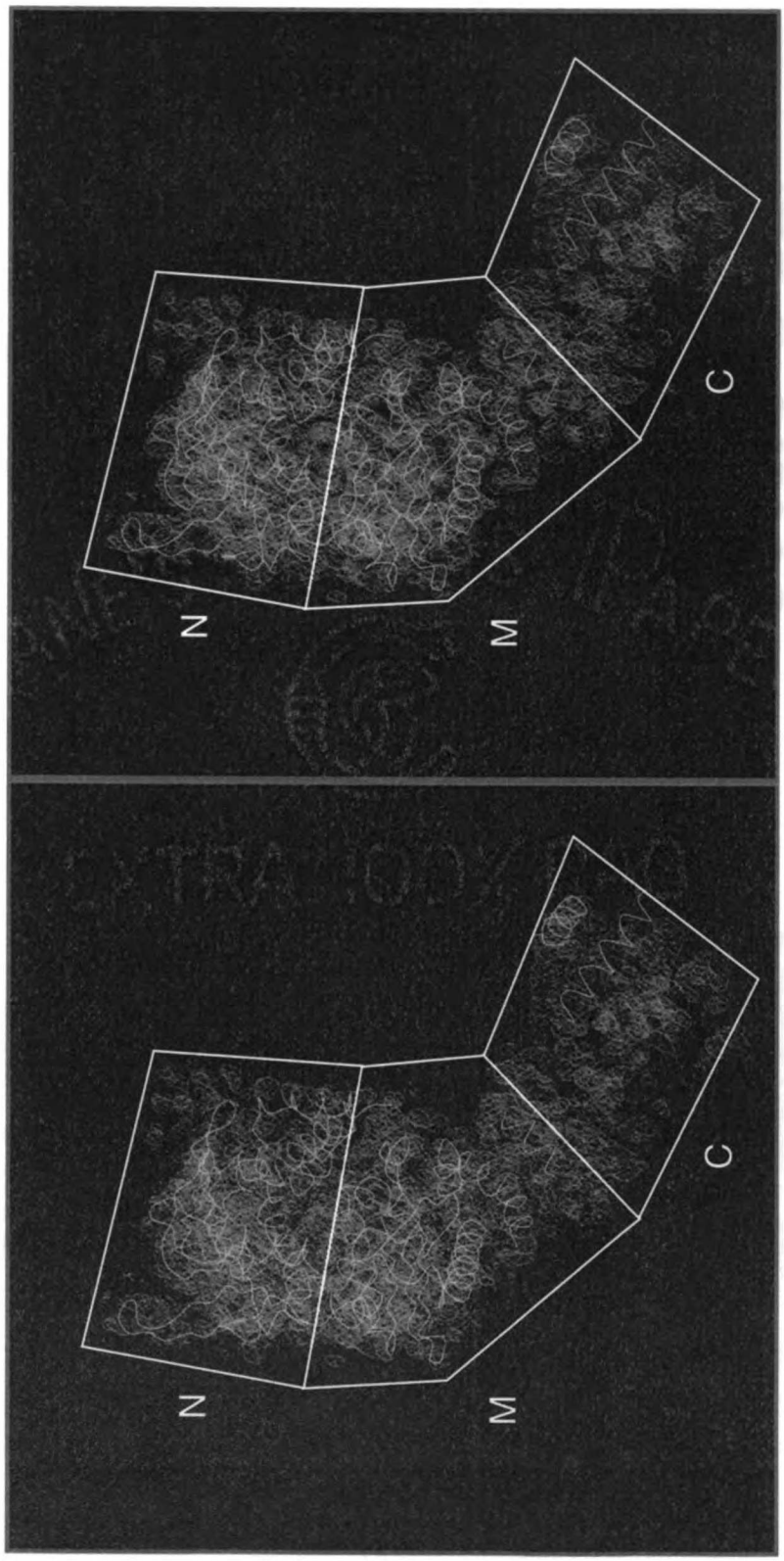
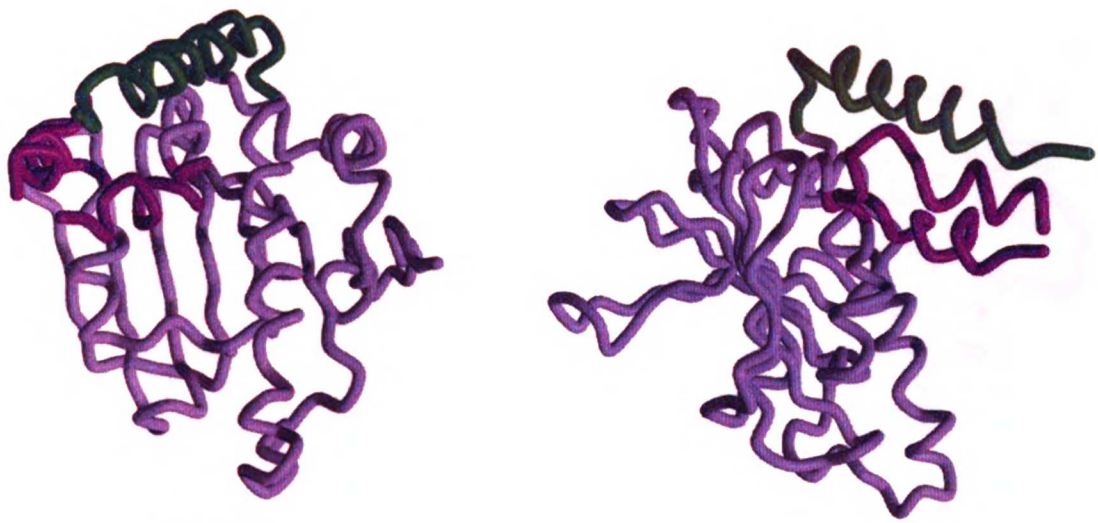
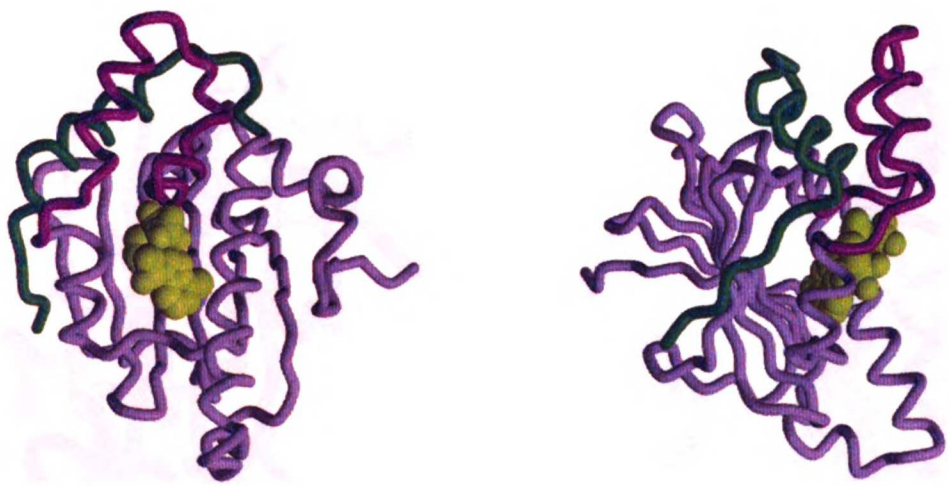


Figure 6

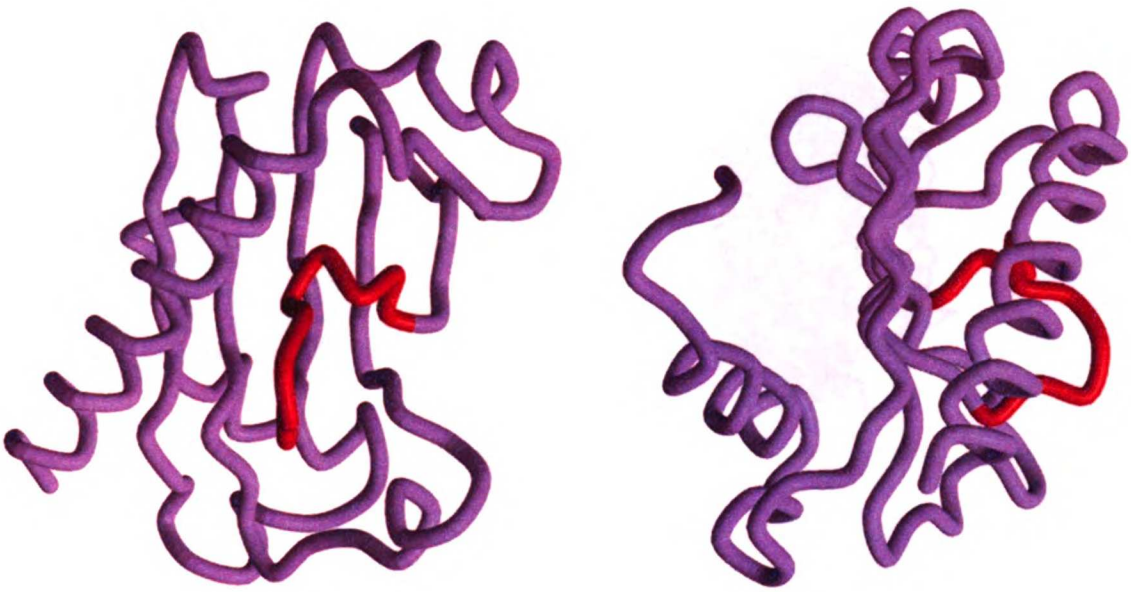


HtpG

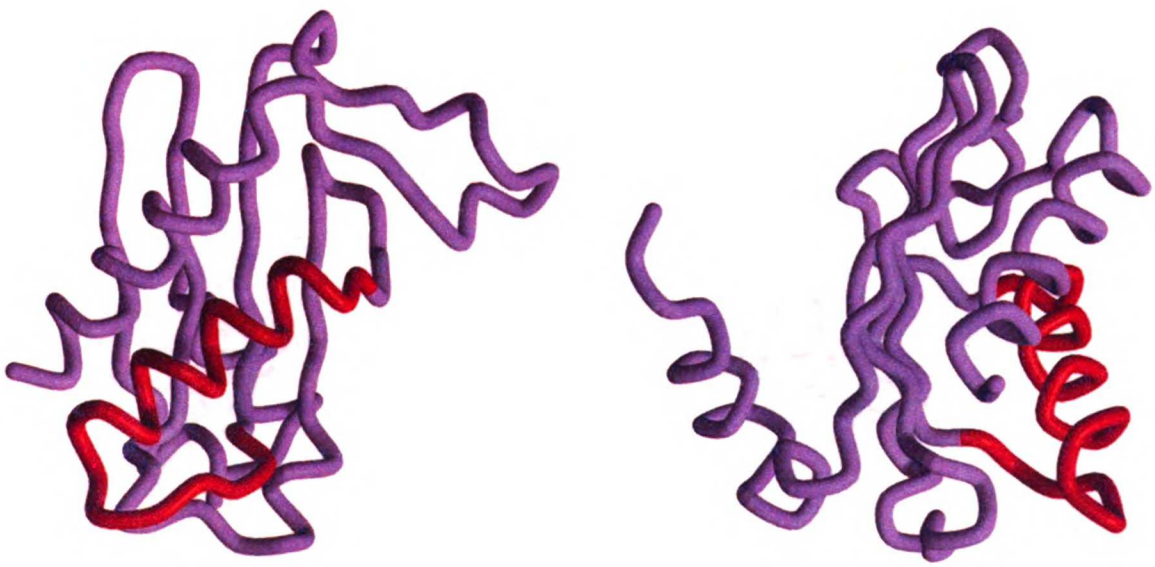


HSP82

Figure 7



HtpG



MutL

Figure 8

Figure 9

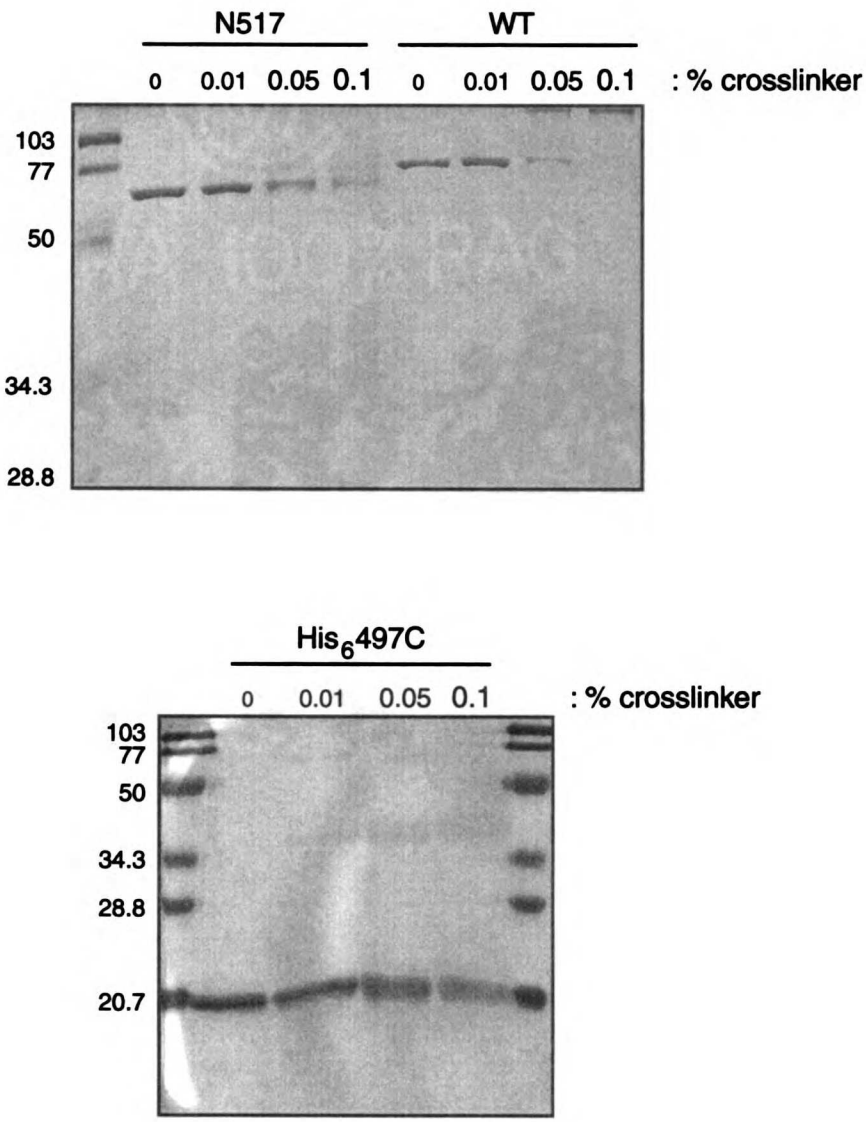

N517
WT
0 0.01 0.05 0.1 0 0.01 0.05 0.1 : % crosslinker
103
77
50
34.3
28.8
His6497C
0 0.01 0.05 0.1 : % crosslinker
103
77
50
34.3
28.8
20.7

Figure 10

Figure 11

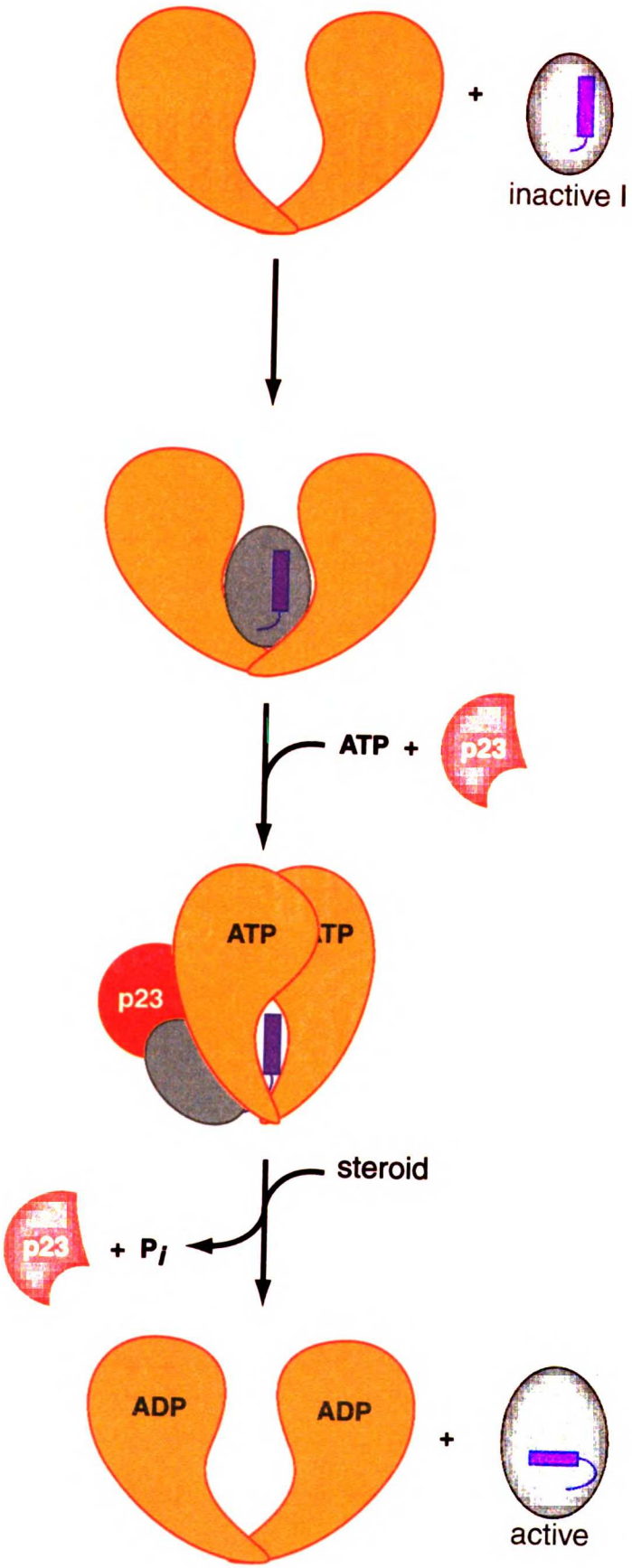


Table 1: HtpG Cysteine Mutants

Mutant	Cloned	Expressed at wild-type levels	Purified	Crystallized (Unmodified)	Crystallized (Derivatized)
S40C	+				
S79C	+	+	+	+	+
S104C	+	+	+	+	
S156C	+	+	+	+	+
S198C	+	+	+		+
T220C	+				
S282C	+	+	+	+	+
A310C	+	+	+	+	+
S335C	+	+	+		
V356C					
S421C					
S437C	+	+	+		
S461C	+	+	+		
S484C	+	+	+		+
S500C	+				
S547C	+				
S590C	+	+	+	+	+
S79C/S156C/A310C	+	+	+	+	+

Table 2: Summary of Crystallographic Statistics (pH 8 Crystal Data)

Data Collection									
Protein	Native	HgS79C	HgS156C	HgS282C	HgA310C	HgS484C	HgS590C	HgS79C/ HgS156C/ HgA310C	SeMet
Crystals	2	1	1	1	1	1	1	1	1
Resolution	4.1	4.1	4.4	4.4	6.0	4.4	5.6	3.9	4.6
Observations	76008	55809	48869	34821	14122	42066	17194	37388	33424
Unique	13793	13803	8057	9468	4688	9150	5813	12012	10617
Completeness (%)	93.4	92.6	66.7	77.4	92.8	74.6	95.4	70.4	98.6
R_{sym} (%)[a]	7.7	13.4	4.8	5.0	7.6	5.8	7.3	10.7	12.5
Average I/σI	10.9	5.9	20.6	14.1	8.4	14.3	9.1	6.4	5.4
MIR Analysis									
Resolution cutoff		4.3	4.4	4.4	6.6	4.4	5.9	4.1	4.6
Sites		2	2	2	2	2	2	6	18
Phasing power[b]		1.50	0.69	0.52	1.74	0.84	0.70	1.09	1.02
R_{Cullis}[c]		0.76	0.92	0.96	0.70	0.91	0.92	0.84	0.87

Overall figure of merit to 4.1 Å = 0.54

[a] $R_{sym} = \Sigma_i |I_i - \langle I_i \rangle| / \Sigma_i I_i$, where $\langle I_i \rangle$ is the average intensity over symmetry equivalents.

[b] Phasing power = $\langle F_h \rangle / \langle E \rangle$, where $\langle F_h \rangle$ is the rms of the heavy atom structure factor amplitude and $\langle E \rangle$ is the rms lack of closure error.

[c] $R_{Cullis} = \Sigma |F_{ph} \pm F_p| - |F_h(calc)| / \Sigma |F_{ph} \pm F_p|$

Table 3: Summary of Crystallographic Statistics (pH 9 Crystal Data)

Data Collection				
Protein	Native	HgS79C	HgA310C	HgS79C/ HgS156C/ HgA310C
Crystals	1	2	1	1
Resolution	3.7	3.9	4.0	4.4
Observations	43512	78166	35254	32563
Unique	17397	16109	14713	8979
Completeness (%)	88.5	94.7	92.8	74.7
R_{sym} (%)[a]	6.2	11.8	9.1	7.7
Average I/σI	10.7	6.8	5.7	7.8
MIR Analysis				
Resolution cutoff		3.9	4.2	4.4
Sites		2	2	6
Phasing power[b]		0.92	0.93	1.26
R_{Cullis}[c]		0.88	0.88	0.78
Overall figure of merit to 3.9 Å = 0.40				

[a] $R_{sym} = \Sigma_i |I_i - \langle I_i \rangle| / \Sigma_i I_i$, where $\langle I_i \rangle$ is the average intensity over symmetry equivalents.

[b] Phasing power = $\langle F_h \rangle / \langle E \rangle$, where $\langle F_h \rangle$ is the rms of the heavy atom structure factor amplitude and $\langle E \rangle$ is the rms lack of closure error.

[c] $R_{Cullis} = \Sigma |F_{ph} \pm F_p| - |F_h(calc)| / \Sigma |F_{ph} \pm F_p|$

Appendix 1

Copyright Permissions

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

ANDREW K. SHIAU
DEPARTMENT OF BIOCHEMISTRY AND BIOPHYSICS
SAN FRANCISCO, CA 94143-0448

TEL: (415) 476-5148
FAX: (415) 476-0939
EMAIL: SHIAU@BIOSYNTH.UCSF.EDU

January 5, 2000

Cell
Fax: (617) 661-7061

To the Editor:

I would like permission to include in my Ph.D. dissertation a copy of the paper cited below, which has been published in Cell. The dissertation will be microfilmed by University Microfilms Incorporated, and they request permission to supply single copies on demand.

Shiau, A. K., Barstad, D., Loria, P. M., Cheng, L., Kushner, P. J., Agard D. A., and Greene, G. L. (1998). The Structural Basis of Estrogen Receptor/Coactivator Recognition and the Antagonism of This Interaction by Tamoxifen. Cell, Vol. 95, 927-937

If possible, please return this letter by fax with the appropriate verification. If there are any problems, please do not hesitate to contact me. Thanks in advance.

Sincerely,

Andrew Shiau