Crystallographic comparison of the estrogen and progesterone receptor's ligand binding domains

DAVID M. TANENBAUM*, YONG WANG*, SHAWN P. WILLIAMS, AND PAUL B. SIGLER†

Department of Molecular Biophysics and Biochemistry and the Howard Hughes Medical Institute, Yale University, 260 Whitney Avenue, JWG 423, New Haven, CT 06511

Contributed by Paul B. Sigler, February 18, 1998

ABSTRACT The 2.8-Å crystal structure of the complex formed by estradiol and the human estrogen receptor-α ligand binding domain (hERαLBD) is described and compared with the recently reported structure of the progesterone complex of the human progesterone receptor ligand binding domain, as well as with similar structures of steroid/nuclear receptor LBDs solved elsewhere. The hormone-bound hERαLBD forms a distinctively different and probably more physiologically important dimer interface than its progesterone counterpart. A comparison of the specificity determinants of hormone binding reveals a common structural theme of mutually supported van der Waals and hydrogen-bonded interactions involving highly conserved residues. The previously suggested mechanism by which the estrogen receptor distinguishes estradiol's unique 3-hydroxy group from the 3-keto function of most other steroids is now described in atomic detail. Mapping of mutagenesis results points to a coactivator-binding surface that includes the region around the “signature sequence” as well as helix 12, where the ligand-dependent conformation of the activation function 2 core is similar in all previously solved steroid/nuclear receptor LBDs. A peculiar crystal packing event displaces helix 12 in the hERαLBD reported here, suggesting a higher degree of dynamic variability than expected for this critical substructure.

Steroid and nuclear hormones exert their influence on target cells by binding to cognate members of the steroid/nuclear receptor superfamily of transcription factors (Fig. 1A). Members of this family are modular in structure with domains A-F, including discrete DNA binding domains and ligand binding domains (LBDs) (3, 4). The receptors are targeted to their respective promoters through specific interaction with cognate hormone response elements (3, 4), and regulate transcription in a ligand-dependent manner with the aid of various coactivators and corepressors (5–8). Additional levels of regulation are achieved through interactions with other systems such as molecular chaperones (4), CAMP-regulated kinases (9, 10), and AP1 activators (4, 11, 12). Members of this superfamily also regulate each other’s expression through complex crosstalk pathways (13, 14). Together these form a rich and elaborate network by which lipophilic ligands regulate gene expression.

The transcriptional response to hormones or antihormones is rooted in conformational changes induced by specifically bound ligands. These changes have been implicated by biochemical and genetic analysis to be involved in transactivation, dimerization, phosphorylation, chaperone interaction, and corepressor inhibition (4). To better understand the molecular mechanism of these responses in chemical terms, crystallographic analysis has been carried out elsewhere on the liganded LBDs of two nuclear receptors, the all-trans retinoic acid receptor (RAR) (15) and the thyroid hormone receptor (TR) (16), and on the unliganded 9-cis retinoic acid receptor (RXR) (17). To extend the structure-function analysis to the steroid subfamily we present here a comparison of the 2.8-Å crystal structure of the estradiol complex of the human estrogen receptor α LBD (hERαLBD) and the progesterone complex of the human progesterone receptor LBD (hPRLBD) (18). A parallel study of the estradiol and raloxifene complexes of hERαLBD recently was published (1), the descriptions of which also are used in the comparison.

Materials and Methods

A fragment of the human ERα (residues 297–554) that contains the entire ligand binding (“E”) domain with an N-terminal three residue (MDP) extension was overexpressed with pET23d in BL21(DE3)pLys S and purified in 5 M urea by estradiol-affinity chromatography (19) with 10 mM ammonium chloride added to prevent carbamylation of lysine residues. The purified protein [a complex of hERαLBD with estradiol in 25 mM Tris (pH 7.4), 100 mM NaCl, 1 mM EDTA, 1 mM DTT, 5 M urea, 10 mM NH₄Cl, 20 μM estradiol] was exchanged into 25 mM Tris (pH 7.4), 200 mM NaCl, 1 mM EDTA, 1 mM DTT, 20 μM estradiol, 0.1% β-octyl glucoside via multiple dialysis steps and concentrated by ultrafiltration to 18 mg/ml. Crystals were obtained by vapor diffusion at 18°C using hanging drops containing an 8-μl 1:1 mixture of the above protein stock solution and a well solution composed of 100 mM Tris (pH 7.6), 480 mM MgCl₂, 10 mM Mg acetate, 10% ethylene glycol, 5% polyethylene glycol (PEG) 4000. Triangular prisms (space group R32; a = b = 147 Å, c = 338 Å in the hexagonal setting; two dimers per asymmetric unit) grew to full size (150 × 150 × 300 microns) in 14 days and diffracted to 3.2 Å. When this same crystal form was stabilized in 0.1 to 10 mM potassium aurolcyanide, a new unit cell formed that had the same space group and a and b axes, but had a halved c axis of 169 Å. Crystals were stabilized in 100 mM Tris (pH 7.6), 510 mM MgCl₂, 10 mM Mg acetate, 10% ethylene glycol, 5.5% PEG 4000, 10 mM KAu(CN)₄ for 18 hr at 18°C; cryoprotected by a 15-sec equilibration in 100 mM Tris (pH 7.6), 510 mM MgCl₂, 10 mM Mg acetate, 25.5% ethylene glycol, and 5.5% PEG 4000, and flash-frozen in liquid propane at liquid nitrogen temperatures. X-ray data were collected at 90–110 K in a nitrogen gas stream. Anomalous diffraction data were collected at NSLS beamline X4a on image plates at the National Synchrotron Light Source, Brookhaven National Laboratory, Upton, NY 11973 (reference nos. 1A52 and R1A52SF).

Abbreviations: LBD, ligand binding domain; PR, progesterone receptor; ER, estrogen receptor; hERαLBD, human estrogen receptor-α LBD; hPRLBD, human progesterone receptor LBD; RAR, retinoic acid receptor; TR, thyroid hormone receptor; RXR, retinoic acid receptor.

Data deposition: The atomic coordinates and structure factors reported in this paper have been deposited in the Protein Data Bank, Biology Department, Brookhaven National Laboratory, Upton, NY 11973 (reference nos. 1A52 and R1A52SF).

* D.M.T. and Y.W. contributed equally to this work.
† To whom reprint requests should be addressed.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.
gold LIII edge. Initial phases were obtained by single wavelength anomalous scattering (SAS) based on gold sites determined from an anomalous difference Patterson synthesis. Using the gold adduct as the parent, phases were improved by multiple isomorphous replacement (MIR) with derivatives produced in the presence of 0.1–1 mM KAu(CN)2 for 12 hr followed by 6 hr in a stabilizer with KAu(CN)2 and a second heavy metal (0.1 mM Au for ethyl mercury chloride, 1 mM Au for mersalyl and trimethyl lead acetate). Data from CHESS-A1 were collected on a 1 K × 1 K charge-coupled device at 0.91 Å. Home-source data were collected by using a Macscience DIP2000 image plate detector and mirror-conditioned Cu Ka. All data were processed by DENZO (20) and merged with SCALEPACK (20). Mercury and lead sites were found by difference Fourier synthesis, and heavy atom parameters were refined with ML-PHARE (21). Combined SAS and MIR phases were further improved by solvent flattening (22) using PHASE (G. Van Duyne, personal communication) and by averaging density related by the noncrystallographic dyad with RAVE (23). The model was built with O (24) and refined by CNS (25) by using a maximum likelihood target, bulk solvent correction, simulated annealing, noncrystallographic symmetry constraint and restraint, and restrained individual B factors. The structure was refined to an R factor of 22.3% and free R factor of 27.4%. See Table 1 for details on data collection and refinement.

**Overall Architecture**

There are two ERαLBDs in the crystallographic asymmetric unit related by a nearly exact noncrystallographic dyad. This arrangement, which resembles that of unliganded RXR,
likely represents the physiologically relevant homodimeric form of ERα in estrogen-dependent transcriptional activation (26), and corresponds to the stoichiometry and rotational symmetry of the ER’s DNA target (27). The fold of the individual ERαLBD is roughly similar to that seen in the structures of other liganded LBDs in the steroid/nuclear superfamily (1.2-Å rms deviation for ER vs. PR for helices 1–11). However, helix 12, which is not essential for the ligand binding or dimerization properties of the ERαLBD (28), extends away from the body of the domain and binds to the surface of a neighboring molecule (Figs. 1B and 2A). This intermolecular contact is somewhat similar to the intramo-

Table 1. Crystallographic analysis

<table>
<thead>
<tr>
<th>Data collection</th>
<th>Parent</th>
<th>Mersalyl</th>
<th>MesPbAC</th>
<th>EtHgCl</th>
</tr>
</thead>
<tbody>
<tr>
<td>X-ray source</td>
<td>NSLS X4a</td>
<td>Laboratory</td>
<td>CHESS-A1</td>
<td>CHESS-A1</td>
</tr>
<tr>
<td>Wavelength</td>
<td>1.0397 Å</td>
<td>1.54 Å</td>
<td>0.91 Å</td>
<td>0.91 Å</td>
</tr>
<tr>
<td>Resolution</td>
<td>50-2.8 Å</td>
<td>40-3.5 Å</td>
<td>40-3.0 Å</td>
<td>40-4.0 Å</td>
</tr>
<tr>
<td>Unique refl</td>
<td>33.401</td>
<td>8.827</td>
<td>14.560</td>
<td>6.029</td>
</tr>
<tr>
<td>Completeness</td>
<td>99.1%</td>
<td>97.9%</td>
<td>99.3%</td>
<td>97.1%</td>
</tr>
<tr>
<td>Rsym</td>
<td>6.8%</td>
<td>10.5%</td>
<td>7%</td>
<td>6.1%</td>
</tr>
<tr>
<td>Reagent conc.*</td>
<td>1 mM</td>
<td>2.5 mM</td>
<td>1 mM</td>
<td></td>
</tr>
<tr>
<td>Soaking time*</td>
<td>6 h</td>
<td>6 h</td>
<td>5 h</td>
<td></td>
</tr>
<tr>
<td>Number of sites</td>
<td>4</td>
<td>8</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>Phasing power†</td>
<td>1.45</td>
<td>1.11</td>
<td>1.1</td>
<td></td>
</tr>
<tr>
<td>Anomalous phasing power†</td>
<td>0.5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Mean f.o.m. 0.45 acentric/0.59 centric

Refinement 50-2.8 Å (2σ data)

Working R factor: 22.3%  Free R factor: 27.4%

Rms deviations  Bond lengths 0.014 Å  Bond angles 1.68°

Rsym = \( \frac{\sum I_h - (I_\delta) / \sum I_h}{} \), where \( I_\delta \) is the average intensity over Friedel and symmetry equivalents. Phasing power = \( \frac{\sum |F_H| - |F_{PHobs}|}{\sum |F_{PHcalc}|} \). Anomalous phasing power = \( \frac{\sum |F_H| - |ADobs|}{\sum |ADcalc|} \), where AD is the Bijvoet difference. NSLS is the National Synchrotron Light Source at Brookhaven National Laboratory, and CHESS is the Cornell High Energy Synchrotron Source.

*See the Materials and Methods section for full details.
†Acentric.

Fig. 2. Arrangement of hERαLBDs in the crystal structure. (A) The tetramer formed by the intermolecular disulfide bonds between Cys-530 of neighboring hERαLBD molecules and dimerization around the local dyad. (Inset) A schematization of the arrangement. Positions of the disulfides are marked with a yellow disc with red “S.” Red symbols in the center of the tetramer and schematic show the crystallographic dyad, and the red arrow in the inset indicates the noncrystallographic symmetry dyad. (B) Comparison of the dimers of holo-hERαLBD, apo-hRXRαLBD, and holo-hPRLBD viewed down the local dyad (17, 18). The dimer interfaces of holo-hERαLBD and apo-hRXRαLBD are similar, with helices 7–10 as the major contributors. The holo-hPRLBD dimer interface is substantially different, composed predominantly of helices 11 and 12, as well as the extreme C-terminal tail. Drawn with RIBBONS (29).
molecular binding mode of helix 12 in the raloxifene-bound hERαLBD (1), where the antiestrogen displaces helix 12 from its position in the estradiol-bound complex. This artificial interaction is caused by the perturbing effect of a fortuitous disulfide bond formed between the loops just preceding helix 12 in neighboring molecules as shown in Fig. 2A. The aberrant position of helix 12 seen in our hERαLBD structure underscores the high degree of positional variabil-

FIG. 3. Specificity determinants of the hormone-binding site specifies 3-hydroxy vs. 3-keto steroids. (A) The hydrogen-bonding network as seen in the hERαLBD and hPRLBD (18). The discriminating relationship of Glu/Gln to the 3-hydroxy/keto of the steroid is supported by a network of water-mediated hydrogen bonds involving the side chain of a conserved Arg and backbone carbonyl of a conserved Phe that, in turn, are fixed by hydrophobic contacts with the steroid ring. Note that the PR has no obvious hydrogen bonding discrimination at the 20-keto position of progesterone comparable to the hydrogen-bonding interaction seen between the 17-hydroxyl of estradiol and His-524. (B) Space-filling representation of estradiol in the ligand-binding pocket of hERαLBD and progesterone in the ligand-binding pocket of hPRLBD (18). For the proteins, carbon atoms are gray, oxygen atoms red, sulfur atoms yellow and nitrogen atoms blue. For the hormones, carbon atoms are cyan, oxygen atoms magenta. Drawn with MIDAS (36). (C) Schematic of estradiol in the hERαLBD ligand-binding pocket in the structure shown here; hormone (red) rings are lettered, and carbon atoms are numbered. Residues hydrogen bonded directly to the hormone are blue. Dashed lines indicate hydrophobic van der Waals contacts with the hormone. Residues conserved among the steroid receptors are green, and variable residues are black. Residues contributed by helix 12 (1) are not shown.

FIG. 4. A stereo view of the 2Fo-Fc Sigma A weighted (37) electron-density map contoured at 1.2σ showing estradiol in the hERαLBD hormone-binding pocket. Drawn with SETOR (38).
receptors such as RAR, TR, and 1,25 dihydroxy vitamin D3 repeat sequences in the ER and TR LBDs (28). Nuclear coil predicted on the basis of the leucine-containing heptad (2). This is distinctly different than the left-handed coiled-coil which maintains an interhelical contact similar to that in resolvase (31), but with a smaller crossing angle (40° for resolvase versus ~29° for ER). This is distinctly different than the left-handed coiled-coil which maintains an interhelical contact similar to that in resolvase (31), but with a smaller crossing angle (40° for resolvase versus ~29° for ER). This domain is oriented as in Fig. 1B. Drawn with RIBBONS (29).

Dimer Interface

Fig. 2B shows the mode of hERα LBD dimerization and its similarity to that seen in the crystal structure of the unliganded hRXRα LBD. The dimer interface observed here has been shown by mutagenesis to be essential for ER dimer formation (28, 30). In the hERα LBD, each protomer contributes 1,477 Å², or roughly 13.3% of its solvent-accessible surface area to the interface, which is comparable to the 1,333 Å², or roughly 11.2% of solvent-accessible surface contributed to the dimer interface of apo-hRXRα (17). Both of these interfaces are larger than that seen in the hPRLBD, which buries 705 Å² of surface for each protomer (18). The much smaller dimer interface of the hPRLBD (18) is distinctly different (Fig. 2B), as the α-helix formed by the C-terminal strand of the hPRLBD intrudes on the homodimer interface seen in the LBD crystal structures of hERα and hRXRα. Gel filtration experiments on a slightly longer fragment of hERα LBD (297–566) confirm that the estradiol-bound hERα LBD is a dimer in solution, whereas the progesterone-bound hPRLBD used for the crystal structure elutes as a monomer (data not shown).

The hERα LBD dimer interface contains residues from helices 7, 8, and 9 as well as the loop between helices 8 and 9, and is dominated by a conserved hydrophobic region at the N terminus of helix 10/11 that forms a right-handed convex 40° for resolvase versus ~29° for ER). This is distinctly different than the left-handed coiled-coil predicted on the basis of the leucine-containing heptad repeat sequences in the ER and TR LBDs (28). Nuclear receptors such as RAR, TR, and 1,25 dihydroxy vitamin D3 receptor are likely to form heterodimers with ER/RXR-like dimer interfaces because the residues that correspond to those in the dimer interfaces seen in the hERα and hRXRα crystal structures are highly conserved in these receptors and have been implicated in dimerization by mutagenesis (17, 32–35).

Hormone Specificity

In the hERα and hPR ligand-binding pockets, the hormone is oriented by two types of contacts: hydrogen bonding at the two ends and hydrophobic van der Waals contacts along the body of the hormone. Estradiol and its analogues are unique in having a 3-hydroxy group as opposed to the 3-keto group found in other steroids. The uniqueness of this estrogenic functionality is reflected in the hydrogen-bonding pattern of the hERα and hPR ligand-binding pockets shown in Fig. 3A and ref. 1. At the end of the binding pocket that holds the A-ring of the steroid, the estrogen receptor is unique in having a glutamate (Glu-353) to accept the hydrogen bond donated by the estrogenic 3-hydroxyl group. The other steroid receptors, which bind 3-keto steroids, have a conserved glutamine at the corresponding sequence position where the hPRLBD structure shows the amido NH₂ group of the Gln-725 donating a hydrogen bond to the 3-keto group of progesterone (18). Fig. 3A also shows how the conserved arginine (Arg-394 in hERα, Arg-766 in hPR) serves to correctly orient and position the discriminating glutamate or glutamine side chain via highly polarized, water-mediated hydrogen bonds. Through a nearly identical structural arrangement in the hERα LBD and hPRLBD, the side chain of this arginine is itself braced by a hydrogen bond to the carbonyl of the residue preceding a conserved phenylalanine (Phe-404 in hERα, Phe-778 in hPR), which is fixed by its van der Waals contact to the A-ring of the steroid (Fig. 3B and C). Thus, the recognition of the 3-oxy function of the hormone is coordinated with a van der Waals contact to its hydrophobic ring. At the other end of the ligand, however, the mechanism of discrimination is less clear. In the hERα LBD, the 17-hydroxyl is hydrogen-bonded to the δ nitrogen of His-524 (Fig. 4; ref. 1), whereas the hPRLBD (18) possesses a less definite discriminating functionality for progesterone’s 17-methyl-keto moiety (Fig. 3A). The hormone is positioned in the same orientation in the hERα LBD as in the hPRLBD (18), but both are in the opposite direction to that predicted on the basis of the nuclear receptor structures (2).

The hydrophobic residues implicated in direct hormone binding in the ER and PR structures are highly conserved (Figs. 1A and 3C), indicating a commonality of pocket architecture. Eighteen residues in ER contact bound estradiol; of those 18 sequence positions 14 are occupied by PR residues that contact bound progesterone and 10 of those have similar side chains in all of the steroid receptors (Fig. 3C). The variations in the van der Waals surface of the binding pocket, however, appear designed to complement the distinctive contours of the cognate hormone. A flat aromatic A-ring, and the absence of a 19-methyl group protruding from its surface distinguish the van der Waals surface of the estradiol molecule. Fig. 3B shows that the bulky side chain of Leu-387 in the ER’s binding pocket is replaced by the smaller, more flexible methionine side chain in PR to accommodate progesterone’s 19-methyl group. These variations on a common theme indicate that it is highly unlikely that the natural ligands of the steroid receptor subfamily bind in a manner that is substantially different than that seen in the ER and PR.

The cooperative stereochemistry of ligand binding provided by the network of specific hydrogen bonds and van der Waals contacts underscores the importance of ligand-induced stability in the steroid/nuclear receptor superfamly. Note that the absence of hormone from the hERα and hPR LBDs would leave cavities of 450 Å³ (1) and 603 Å³ (18), respectively and thus would require significant readjustments to retain stability of the hydrophobic interior of the domain. In the two hormone-bound LBDs discussed here, it is clear that this cooperative network of hormone contacts not only accounts for the specificity of ligand binding but allows the hormone to function as the structural core for the bottom half of the LBD, and therefore the scaffold around which this region folds.

Coactivator Binding

Coactivators are necessary for the hormone-bound ER and PR to stimulate transcription (39). The ligand-dependent interac-
tion between steroid/nuclear receptors and coactivators depends on the activation function 2 (AF-2) core located in helix 12 (40). Crystal structures comparing the unliganded LBD from RXR and its liganded counterpart from RAR indicate that the bound hormone induces a dramatic repositioning of helix 12 (15, 17). Except for the artefactual alignment of helix 12 in the ER complex reported here, the structures of all hormone-bound LBDs solved to date (RAR, TR, ER, and PR) shows helix 12, and hence the AF-2 core in the same position (1, 15–18). Importantly, the crystal structure of the antiestrogen, raloxifene, bound to the hERαLBD (1), as well as modeling of the antiprogestin, RU486, bound to the hPR-LBD (18), shows a displacement of helix 12 from this common ‘active’ position, thereby reinforcing the generally accepted view that the active state produced by the hormone is caused by the contribution of the conserved helix 12 to a coactivator-binding surface. Helix 12 is fixed in this active position both by contacts with the hormone and residues on the surface of the domain (Fig. 5). The stabilizing effect originally attributed to an ionic interaction in RARγ between the conserved glutamate of the AF-2 core and a lysine in helix 4 does not seem to be general (15). In TRα, ERα, and PR, stability appears to arise from hydrophobic van der Waals contact between the invariant hydrophobic side chain that follows the conserved glutamate of helix 12 and the hydrophobic components of the side chains from helix 4 (1, 16, 18). In addition to helix 12, residues in ERα such as Lys-362 (41) and Val-364 (42) in the highly conserved loop between helices 3 and 4 (the so-called “signature sequence” (2) indicated in Fig. 1A) have been implicated by mutagenesis as contributing to a substructure responsible for coactivator interaction. These results suggest a coactivator-binding interface that appears to encompass the exposed surfaces of helices 3, 4, and 12, and the “signature sequence” loop, as seen in Fig. 5.

Hormone-dependent activation is predicated on the movements of helix 12 from one position to another. It is noteworthy in this regard that helix 12 as seen in the unliganded RXR, the raloxifene complex of ERα, and the estrogen receptor complex (CCP4), respectively. D.M.T. was a National Institutes of Health predoctoral fellowship student. This work was supported in part by National Institutes of Health Grant GM15225.

35. Computer Graphics Laboratory, University of California, San Francisco (1989) MADSPLUS (Univ. of California, San Francisco).