ELSEVIER

The conserved residues of the ligand-binding domains of steroid receptors are located in the core of the molecules

Miki Maeda

Biochemistry Department, National Institute of Agrobiological Sciences, Tsukuba, Ibaraki 305-8602, Japan

The relationship between conserved residues and biochemical functions of steroid receptors was investigated. Pairwise three-dimensional (3D) alignment of the ligand-binding domains of the human estrogen (1A52) and progesterone (1A28) receptors revealed two conserved domains; Asn313-Ser456 and Gln471-Lys531 (numbering reflects the sequence in the human estrogen receptor). Alignment of the protein sequences of 39 steroid receptors revealed 36 highly conserved residues (i.e., the residues commonly found in more than 80% of sequences aligned). They were distributed throughout the sequences but formed a contiguous 3D structure. Most of these highly conserved residues were buried in the ligand-binding domain, but several residues were exposed on the surface. The well-known functions commonly associated with the ligand-binding domain of steroid receptors are ligand binding, HSP90 binding, transcriptional activation and dimerization. The relationship between the residues and these functions were checked. To determine the residues involved in dimerization, the differences between the solvent accessibilities of the monomeric and dimeric forms were calculated. These results revealed 32 residues of 1A52 and 15 residues of 1A28 potentially involved in dimerization. Their distribution areas do not overlap greatly. Comparing these putative dimerization sites with highly conserved residues, many of the exposed conserved residues were observed on the side of the domain opposite are the dimerization sites. Some highly conserved residues are located in a steroid-binding site and in transcriptional activation domain. However, few of them were observed in the HSP90 binding site. These results indicate that the core structure made by most of the highly conserved residues among the ligand-binding domains of steroid receptors is important. These conserved residues may be essential for conformational change in the ligand-binding domain from its inactive to active form. © 2001 by Elsevier Science Inc.

Keywords: Ligand-binding domain, steroid receptors, 3D alignment, solvent accessibility, molecular dimerization, estrogen receptor, progesterone receptor

INTRODUCTION

Steroid receptors belong to the steroid/retinoid receptor superfamily and share a common structure that is separated into the modulating, DNA-binding, and ligand-binding domains[1–5]. The sequence homology among the ligand-binding domains is lower than that among the DNA-binding domains, and each domain type has its own independent mutation rate depending on its function[6,7]. The degree of homology at a position may relate to the importance of that residue.

The ligand-binding domain of steroid receptors has several biochemical roles: binding to hormone, dimerization, and binding to another protein such as HSP90 (90 kDa heat-shock protein) or the p160 family of coactivators (160 kDa transcriptional coactivator). Although steroid receptors share some sequence homology in the ligand-binding domains, a correlation between the conserved residues and the sequences involved in the various domain-associated functions is still unclear. Steroid receptors make homodimers their active forms, whereas the other nucleic receptors form heterodimers with retinoid X receptor.

In addition, each androgen, estrogen, and gestagen has a unique steroid receptor, whereas corticoids can bind to mineralocorticoid as well as to glucocorticoid receptors; the biochemical response to corticoids varies depending on which particular hormone-receptor complex is formed. The receptors recognize differences between ligands. The ligand-binding domains of steroid receptors are very interesting in regard to

Color Plates for this article are on pages 601 and 606.

Corresponding author: Biochemistry Department, National Institute of Agrobiological Sciences, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan Address: Biochemistry Department, National Institute of Agrobiological Sciences, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan Tel.: +81-298-38-7463; fax: +81-298-38-7417

E-mail: mmaeda@nias.affrc.go.jp

Journal of Molecular Graphics and Modelling 19, 543–551, 2001

655 Avenue of the Americas, New York, NY 10010

1093-3263/01/$–see front matter
PII: S1093-3263(01)00087-0

structure-function relationships, especially for the mechanism by which these domains "choose" their own ligand. Clarifying this mechanism could give insight for designing new synthetic ligands.

To determine the importance of the residues, the conservation rate can be applied as a scale. The difference in homology between the DNA-binding and ligand-binding domains seems to reflect the importance of each domain. Residues conserved among a group of proteins with similar activities often occur in the functional domains. PROSITE[9] is a well-known database that contains many of the conserved sites, called motifs. Many of the motifs in the database correspond to sequentially local homologous sites in a group of proteins. The number of reported 3D protein structures is growing rapidly. As the result, we can analyze which motifs are important for specific functions. Besides sequential motifs, structural motifs are expected to play a role in the various functions of the proteins.

To detect the common residues/sequences, multiple alignment is generally performed. It is effective but has several problems including results sensitive to changes in parameter settings and the occasional requirement for manual correction. To eliminate manual estimation and to obtain objective, reproducible results, a scheme to correct the result of basic multiple alignment estimated by 3D alignment was carried out in this study. Ligand-binding domains of human estrogen receptor and human progesterone receptor have almost 40% amino acid sequence homology (about 25% identity), and some gap insertions are needed for the alignment. The position and length of the gaps depend on the gap penalty scores set by the user. Using the 3D alignment of two proteins, a more objective result from multiple alignment is obtained.

To improve our understanding of the mechanisms of ligand recognition and expression of function, it is necessary to clarify the relationship between structures and various biochemical characteristics. In the functions of steroid receptors, ligand binding and HSP90 binding/8S complex formation are well studied experimentally[10–13]. Because candidate residues for dimerization are not clear, a new method to determine the site of interaction is proposed in this study. The candidates are picked by calculated differences of solvent accessibility between a free and dimer structures. The accessibility value changes when two atoms are closely located to each other. Interaction between two atoms is often estimated from distances between them; in contrast, this method provides quantitative information useful for comparison of aligned structures.

Using these results, the relationship between the conserved residues and biochemical functions are explored and discussed in this paper.

METHODS

Data Set of Proteins

The 37 amino-acid sequences in Table 1 were obtained from SwissProt, release 38.0 (June 1999). They encode the complete ligand-binding domains of the nuclear androgen, estrogen, gestagen (i.e., progesterone), glucocorticoid, and mineralocorticoid receptors from various vertebrates. To obtain the sequence used as the ligand-binding domain of a particular receptor, the complete receptor sequences for all members of that group were aligned, and the residues corresponding to the known ligand-binding domain of the human protein were selected. Insights into the structure-function relationship of the ligand-binding domains were gained by comparing the sequences in Table 1 with sequence and structure information obtained from the complete ligand-binding domains of the human estradiol[10] (1A52) and progesterone[11] (1A28) receptors (Protein Data Bank, release 92). Information about the structure of these two ligand-binding domains arises from X-ray diffraction studies with a resolution of 2.8 Å (1A28) or 1.8 Å (1A52).

Sequence Alignment and Structural Analysis

Simple alignment of the sequences in Table 1 was accomplished by using the PILEUP program in the GCG sequence analysis package version 10[14,15] with default parameters. The sequences of the most highly conserved contiguous region (corresponding to Asn359-Lys401 of 1A52 and Lys731-Gly773 of 1A28) comprised the data set for the initial 3D alignment (cutoff value 3 Å; Protein Adviser, version 3.5, Fujitsu Kyushu System Engineering Limited, Fukuoka, Japan). This initial alignment revealed two regions in 1A52 and 1A28 (Pro324-His398 of 1A52 and Pro696-Val770 of 1A28 and Glu443-Ser456 of 1A52 and Glu816-Glu443 of 1A28) whose structures were very similar. These sequences comprised the data set for a subsequent 3D alignment (cutoff value 3 Å). The final 3D alignment was made by using Asp313-Ser456 of 1A52 and Pro685-Thr829 of 1A28 and Glu471-Lys531 of 1A52 and Ser837-Gln897 of 1A28. Distances between corresponding atoms in excess of 3 Å were measured directly. The computer-generated multiple sequence alignment was adjusted manually on the basis of two sets of sequences (Asp313-Ser456 of 1A52 and Pro685-Thr829 of 1A28 and Glu471-Lys531 of 1A52 and Ser837-Gln897 of 1A28) whose corresponding structures were highly homologous.

In addition, Protein Adviser was used to estimate the location of the conserved residues, which were plotted manually on the simulated structures. The values for the solvent accessibility of highly conserved residues were calculated by using the DSSP program,[16] version Oct. 1988. Solvent accessibility (SA) was calculated and expressed as the number of accessible water molecules in contact with the residues as defined by the DSSP program. Residues whose solvent accessibility differed by at least 0.10 between the monomer and dimer forms were defined as candidates for dimerization.

RESULTS AND DISCUSSION

Previously known are four main biochemical roles of ligand-binding domains of steroid receptors: (1) to bind steroids, (2) to bind HSP90, (3) transcriptional activation, and (4) dimerization. The ligand-binding domains that bind HSP90 are observed as non-activated forms. Binding to steroids causes some conformational changes that result in release of HSP90 and dimerization of the receptors themselves. Because the other members of steroid/retinoid receptor superfamily form a heterodimer with the retinoid X receptor,[8] they should have an independent mechanism for dimerization (however, only steroid receptors are considered in this study). The relationship of our alignments with each of the four biochemical functions will be discussed in the following paragraphs.

Table 1. Amino-acid sequences of the ligand-binding domains of various steroid receptors

SwissProt ID	Accession Number	Species	Residues
Androgen receptors			
ANDR_HUMAN	P10275	*Homo sapiens*	690–919
ANDR_MOUSE	P19091	*Mus musculus*	670–899
ANDR_RABIT	P49699	*Oryctolagus cuniculus*	480–709
ANDR_RAT	P15207	*Rattus norvegicus*	673–902
Estrogen receptors			
ESR1_CHICK	P06212	*Gallus gallus*	305–589
ESR1_HUMAN	P03372	*Homo sapiens*	311–551
ESR1_MOUSE	P19785	*Mus musculus*	315–599
ESR1_OREAU	P50240	*Oreochromis aureus*	266–583
ESR1_ORYLA	P50241	*Oryzias latipes*	315–620
ESR1_PIG	Q29040	*Sus scrofa*	311–551
ESR1_PAGMA	O42132	*Pagrus major*	273–581
ESR1_POEGU	Q91250	*Poephila guttata*	303–587
ESR1_RAT	P06211	*Rattus norvegicus*	316–600
ESR1_SALIR	P16058	*Salmo irideus*	277–574
ESR1_SALSA	P50242	*Salmo salar*	237–535
ESR1_XENLA	P81559	*Xenopus laevis*	303–586
ESR2_HUMAN	Q92731	*Homo sapiens*	215–530
ESR2_MOUSE	O08537	*Mus musculus*	215–530
ESR2_RAT	Q62986	*Rattus norvegicus*	215–530
Progesterone receptors			
PRGR_CHICK	P07812	*Gallus gallus*	487–786
PRGR_HUMAN	P06401	*Homo sapiens*	681–933
PRGR_MOUSE	Q00175	*Mus musculus*	671–923
PRGR_RABIT	P06186	*Oryctolagus cuniculus*	678–930
Glucocorticoid receptors			
GCR_AOTNA	P79686	*Aotus nancymaae*	528–777
GCR_CAVPO	P49115	*Cavia porcellus*	482–771
GCR_HUMAN	P04150	*Homo sapiens*	528–777
GCR_MOUSE	P06537	*Mus musculus*	494–783
GCR_ONCMY	P49843	*Oncorhynchus mykiss*	462–758
GCR_PAROL	O73673	*Paralichthys olivaceus*	520–807
GCR_RAT	P06536	*Rattus norvegicus*	506–795
GCR_SAGOE	P79269	*Saguinus oedipus*	528–777
GCR_SAIBB	O13186	*Saimiri boliviensis boliviensis*	528–777
GCR_SAISC	O46567	*Saimiri sciureus*	528–778
GCR_TUPGB	Q95267	*Tupaia glis belangeri*	486–776
GCR_XENLA	P49844	*Xenopus laevis*	486–776
Mineralocorticoid receptors			
MCR_HUMAN	P08235	*Homo sapiens*	733–984
MCR_RAT	P22199	*Rattus norvegicus*	730–981
MCR_TUPGB	Q29131	*Tupaia glis belangeri*	726–977
MCR_XENLA	Q91573	*Xenopus laevis*	361–612

Identification of Conserved Residues

A total of 204 residues of the ligand-binding domains of the estrogen receptor (1A52) and progesterone receptor (1A28) were aligned and yielded highly homologous structures; the mean distance between their respective α-carbons is 1.15 Å (Color Plate 1a). The residues shown in Color Plate 1a lie in structurally well conserved regions of the ligand-binding domain (Color Plate 1b), but the structural homology between other portions of the aligned domains varies widely (Color Plate 1c). The C-terminal helix (helix 20) of 1A52 lies in a direction different from that of 1A28, and the direction and length of the internal loops of 1A52 and 1A28 differ.

On the basis of the hypothesis that conserved residues have important roles in protein function, highly conserved residues (i.e., those present in at least 80% of the aligned sequences) were identified (Color Plate 2). This identification revealed that

21 amino acids were conserved perfectly among the 39 aligned sequences, nine residues were identical in more than 35 (90%) of the samples, and six residues were homologous in more than 31 (80%). A total of 36 (14%) of the 262 residues aligned were determined to be highly conserved. Of the 36, 80% (29 residues) were distributed in the C-terminal half but did not make any continuous sequences longer than three residues.

Location of Highly Conserved Residues

To gain insight into the potential role of these 36 highly conserved residues, they were plotted on the structures of 1A52 (Color Plates 3a and b) and 1A28 (Color Plates 3c and d). Most of these residues form a contiguous structural core with the ligand inside the ligand-binding domain, and the core structure was surrounded by non-conserved residues (Color Plate 3 and Color Plate 4). Many of conserved residues are included in α-helices and some of them connect sequentially separated α-helices. This result suggests that the conserved residues are necessary to maintain the 3D structure of the ligand-binding domain. Non-conserved residues may perform roles in receptor-specific functions.

To estimate the exposure of the residues, the solvent accessibility (SA) of each residue was calculated. It is an index for characterizing the local environment of the residues; thus the value can be applied to estimate the difference between the two states. The greater the SA, the more likely the residue is located on the surface of the molecule, and a larger ΔSA means that more differences are observed between the two states. Highly conserved residues can be grouped to three by SA values; (1) ΔSA is relatively high ($>$ 0.4), (2) low ΔSA and high SA ($>$0.4), and (3) low ΔSA and low SA ($\leqq$ 0.4). Each group of the residues will be discussed separately.

Table 2. Residues adjacent to ligands

1A52		1A28	
Number	Residue	Number	Residue
343	Met	715	Leu
346	Leu[a]	718	Leu[a]
		719	Asn
349	Leu[a]	721	Leu[a]
350	Ala	722	Gly
353	Glu[c]	725	Gln[c]
		755	Trp[a]
384	Leu	756	Met
387	Arg	759	Met
388	Met	760	Val
391	Leu[a]	763	Leu[a]
394	Arg[a,c]	766	Arg[a,c]
404	Phe[a,c]	778	Phe[a,c]
421	Met		
424	Ile	797	Leu
428	Leu	801	Met
521	Gly[d]	887	Leu
524	His	890	Tyr
525	Leu	891	Cys
		894	Thr
		903	Val[b]
		905	Phe[b]
		909	Met[b]

[a] Highly conserved residues (i.e., present in more than 80% of sequences studied).
[b] Residues in the C-terminal variable loops.
[c] Residues previously reported as interacting mainly with the ligand.[17,20]
[d] Change of Gly521 to Arg/Asp prevent to bind 17β-estradiol.[14]

Polymorphism of Helix 12 Location in Human Estrogen Receptor

Comparing 3D structures of 1A52 and 1A28, the most obvious change involves the C-terminal helices. The residues of Val903, Phe905, and Met909 in 1A28, which are not conserved residues, were observed near progesterone, but no C-terminal residues of 1A52 were detected near the ligand (Table 2). Another structure of the ligand-binding domain of the estrogen receptor with 17β -estradiol (i.e., 1ERE) has been reported.[17] In the structure of 1ERE, the C-terminal helix occupies a location similar to that in 1A28 (progesterone receptor). This similarity suggests that the C-terminal helix change conformationally between at least two conformers. A comparison of helix 12 of 1A52 and 1ERE is shown in Color Plate 3e. The 3D structures of other estrogen receptors were also analyzed. Visualization of three complexes (i.e., 1QKM estrogen receptor-β with genistein;[18] 3ERD estrogen receptor-α with diethylstilbestrol;[19] and 3ERT estrogen receptor-α with 4-hydroxytamoxifen[19]) shows that the C-terminal helices are generally located in a position similar to that in 1ERE. Only the highly conserved residue of Glu542 in 1A52 (corresponding to Glu913 in 1A28) is located in the C-terminal helix 12. As shown in previous report,[20] this helix 12 is important for transcriptional activation.

Conserved Residues with Large ΔSA Values

Consider corresponding pairs of residues in 1A52 and 1A28 with relatively high ($>$ 0.4) ΔSA: Trp383-Trp755, Trp393-Trp765, Leu469-Leu835, and Glu542-Glu911 of 1A52-1A28. Among these conserved residues with high ΔSA values, only glutamate changes its location between 1A52 and 1A28 as discussed in the above section (Color Plates 4c and 4g); this change in conformation leads to the ΔSA between Glu542 and Glu911. The remaining three pairs of conserved residues (Trp383-Trp755, Trp393-Trp765, and Leu469-Leu835) do not vary with respect to their 3D structure; for each pair of these hydrophobic residues, the location of the α-carbon is almost the same. However, the degree of solvent exposure (the solvent accessibility) differs for each component of the pair (Table 3; Color Plates 4d and 4h, 4b and 4f, and 4a and 4e). This result indicates that the packing of the side chains varies slightly. Therefore, these sites might play a role in the functions of the ligand-binding domain, vary depending on the specific steroid receptor.

Conserved Residues on the Surface of the Receptors

Most of the conserved residues are buried in the interior of the proteins, and only several residues are observed on the surface of the receptor (Color Plate 4). The conserved residues on the

Table 3. Solvent accessibility of highly conserved residues

1A52			1A28			
Number	Residue	SA[a]	Number	Residue	SA[a]	ΔSA[b]
319	Leu	0.01	691	Leu	0.01	0
323	Glu	0.58	695	Glu	0.52	0.06
324	Pro	0.41	696	Pro	0.51	0.10
336	Pro	0.93	708	Pro	1.19	0.26
346	Leu	0.22	718	Leu	0.25	0.03
349	Leu	0.05	721	Leu	0.05	0
360	Trp	0.03	732	Trp	0.12	0.09
361	Ala	0	733	Ser	0.03	0.03
362	Lys	1.23	734	Lys	1.44	0.21
365	Pro	0.16	737	Pro	0.43	0.27
366	Gly	0.08	738	Gly	0.29	0.21
367	Phe	0.02	739	Phe	0	0.02
370	Leu	0.10	742	Leu	0.06	0.04
374	Asp	0.17	746	Asp	0.05	0.12
375	Gln	0.26	747	Gln	0.27	0.01
378	Leu	0	750	Leu	0	0
383	Trp	0.57	755	Trp	0.05	0.52
391	Leu	0.10	763	Leu	0.04	0.06
393	Trp	0.72	765	Trp	0.23	0.49
394	Arg	0.37	766	Arg	0.42	0.05
395	Ser	0	767	Ser	0	0
402	Leu	0	776	Leu	0	0
404	Phe	0.29	778	Phe	0.20	0.09
405	Ala	0	779	Ala	0	0
408	Leu	0.04	782	Leu	0	0.04
444	Glu	0.10	817	Glu	0.03	0.07
449	Lys	0.11	822	Lys	0.16	0.05
453	Leu	0	826	Leu	0.01	0.01
454	Leu	0.02	827	Leu	0	0.02
469	Leu	0.78	835	Leu	0.13	0.65
486	Leu	0	852	Leu	0	0
503	Arg	0.35	869	Arg	0.17	0.18
506	Gln	0.76	872	Gln	0.85	0.11
507	Leu	0	873	Leu	0	0
511	Leu	0.20	877	Leu	0	0.20
542	Glu	1.71	911	Glu	1.15	0.56

[a] Solvent accessibility; the number of water molecules in contact with the residues.
[b] ΔSA is calculated as the difference between the residues' SA of 1A52 and 1A28.

surface can be candidates involved in a common function (e.g., interacting to another common molecule) among steroid receptors.

To observe the exposure of conserved residues on the surface, space-filling models were drawn (Color Plate 4). Color Plates 4a and 4e represent the molecular surface of the ligand-binding domains corresponding that in Color Plates 3a and 3b, respectively. Color Plates 4a and 4e display one side of the protein, and Color Plates 4c and 4g show the opposite face. Several conserved residues can be seen on the protein surface of each face. Because more than a half of the conserved residues on the surface are located on the half illustrated in Color Plates 4c and 4g, this region might perform a biochemical role common to steroid receptors.

Consider conserved residues: the five pairs of Glu323-Glu695, Pro324-Pro696, Pro336-Pro708, Lys362-Lys734, Gln506-Gln872 (the residue in 1A52 is first, followed by that of 1A28). These residues have relatively high (> 0.4) SA and low (< 0.3) ΔSA values (Table 3). They can be candidates for common functional sites in steroid receptors. However, for four residue pairs (Trp383-Trp755, Trp393-Trp765, Leu469-Leu835 and Glu542-Glu911), 1A52 has high (>0.4) ΔSA values (Table 3) but 1A28 does not. This suggests that these residues perform different functions between or a structural change is needed to express the common function using these residues.

Potential Meaning of the Residues Close to the Ligands

The most important role of the ligand-binding domain is to bind steroids. For sake of this study, these domains are defined to consist of those residues with non-hydrogen atoms within 5.0 Å of non-hydrogen atoms of ligands (Table 2). The distance of 5.0 Å was selected to include hydrogen bonding residues. The O-O length of hydrogen bond in water is less than 3.0 Å. Five highly conserved residues meeting this criterion occurred in both 1A52 and 1A28 (Leu346, Leu349, Leu391, Arg394, and Phe404 of 1A52), and an additional such residue occurred only in 1A28 (Trp755). The five residues common to both structures are located on one side of the ligand. Arg394 of 1A52 (Arg766 of 1A28) near the oxygen of the A-ring can be estimated to make a hydrogen bond, and the other four conserved residues lying on the opposite side of carbon-18 of the steroid skeleton can be estimated to contact the steroids by hydrophobic interaction. These five highly conserved residues are included in the 18 residues previously reported as the residues contacting estrogen and progesterone.[10,11] The ligand structure recognized by these five residues is likely the common structure among most native steroid ligands. The remaining residue, Trp755 of 1A28, lies close to progesterone, but Trp383 (the corresponding residue in 1A52) is relatively far from estrogen. This difference seems to be caused by absence of the carbon-19 methyl group of the steroid. Because Trp383 of 1A52 and Try755 of 1A28 occupy similar position in the ligand-binding domains (Color Plate 3), these bulky residues likely comprises part of the steric binding pocket.

Conserved Residues Potentially Involved in HSP90 Binding/8S Complex Formation

It is generally recognized that the ligand-binding domain of the steroid receptor is also involved in HSP90 binding. The 8S conformer (the binding form with HSP90) is inactive and becomes the active form upon release from HSP90. The detail mechanism is unclear, but steroid binding and HSP90 binding seem to be related.

Candidates of the interaction sites in human glucocorticoid receptor were proposed.[12,13] Cadepent et al. determined HSP90 binding (i.e., 8S complex formation) and steroid binding activities using a series of deleted receptor derivatives.[12] They reported that the mutant from the 1st to 696th residues was fully active for making the 8S complex but not the steroid complex

and that the mutant consisting of residues 1 to 550 lacks the ability for making the 8S complex and for steroid binding. However, the mutant combining the 1 to 550 sequence with the sequence 627 to 696 is fully active to make the 8S complex. In sequence from residues 408 to 486 of the estrogen receptor (corresponding to the 627th to 696th residues of human glucocorticoid receptor), six highly conserved residues are observed (the corresponding residues in 1A52 and 1A28, respectively, being Leu408-Leu782, Glu444-Glu817, Lys449-Lys822, Leu453-Leu826, Leu454-Leu827 and Leu469-Leu835 of 1A52-1A28). Because the five residues other than Leu469-Leu835 have a SA value less than 0.20, they are assumed to be buried in the molecules. If so, these five molecules would not be involved in the molecular interaction directly. Leu469-Leu835 is located in the internal loop region and Leu469 of human estrogen receptor has a relatively high SA value (0.78). The residue can thus be accessed by another molecule. The 3D structural situation is different in estrogen and progesterone receptors. Leu469 of estrogen receptor and Leu835 of progesterone receptor may play a role in making an 8S conformational change.

Xu et al. reported that the amino acid sequence from the 547th to 553rd of human glucocorticoid receptor (corresponding to the 312th to 318th residues of the human estrogen receptor) are involved in HSP90 binding[13]. Highly conserved residues detected in this study are not in this region. Analyzing the 3D structure around Ala312-Ala318, it was located close to a helix containing Lys481-Leu489 (Color Plate 5). The interacting site contains the conserved residue Leu486 of the human estrogen receptor. The length of the segment Ala312-Ala318 just fits the width of the helix containing Lys481-Leu489. This observation may mean that the sequence of Ala312-Ala318 is essential to maintain the location made by two helices of Leu467-Leu497 and Gln441-Ser456 (in the human estrogen receptor).

Conserved Residues Potentially Involved in Transcriptional Activation

Transcriptional activation is mediated by two activation domains, AF1 at the N terminus and AF2 in the ligand-binding domain.[21–23] Previous reports showed AF-2a was located in the first helix corresponding to the sequence 312 to 318 in the ligand-binding domain of the human estrogen receptor.[24–26] It seems meaningful that AF-2a is close to the HSP90 binding site. These two sites are sequentially continuous and in the same helix. The AF-2a domain contains three highly conserved residues. Pierrat et al. reported that one more aspartate (Asp332 of human estrogen receptor) was also conserved,[24] but the aspartate was not observed in all estrogen receptors-β and estrogen receptors-α of fish. Thus, we neglected the extra aspartate in this study (refer to Color Plate 2). The aspartate would be less important than Leu319, Glu323 and Pro324 (numbering is corresponding to the human estrogen receptor) on transcriptional activation.

Mak et al. pointed out that helix 12 and Lys366 are required for binding the p160 family of coactivators.[27] The p160 (or RIP160)[28,29] family of coactivators appears to bind to most, if not all, nuclear receptors in a ligand-dependent manner and causes transcription of target genes. They thought the surface area involved in the binding to a transcriptional coactivator was made by Ile362, Lys366, Leu376, Val380, Leu543 and Leu546 of mouse estrogen receptor (corresponding to Ile358, Lys362, Val376, Leu539 and Leu542 of human estrogen receptor, respectively). The authors showed that Leu543 (corresponding to Leu539), highly conserved among human proteins, is important; the change Leu543 to alanine decreased the activity for transcriptional activation and suggested the importance of this residue. However, the amino acid residue corresponding to Leu543 of the mouse estrogen receptor is conserved as methionine among all androgen, progesterone, mineralocorticoid, and glucocorticoid receptors as shown in Color Plate 2. Mak et al. showed Leu543 is near Glu546.[27] Thus, the change to methionine would retain remaining the interaction to the aspartate. As described above, this aspartate is very conserved and is located in a movable part of the receptor because two crystal structures (1A82 and 1ERE) show different locations of the helix around it. It may suggest that the structural change caused by ligand-binding makes the helix containing Glu546 and Leu543 of the mouse estrogen receptor to shift the location observed in 1ERE and the change is important for binding the p160 family protein.

Phosphorylation site and conserved residues Tyr537 in the human estrogen receptor is a well-known residue as a phosphorylation site. Yudt et al.[30] reported the mutation effect Tyr537 to Phe537. They showed that the phosphorylation is not required for the human estrogen receptor to bind hormone or to activate transcription. They also showed that mutation caused a change in estradiol binding kinetics. The authors proposed that the mutation from Tyr537 to Phe537 favors an "open" pocket conformation, affecting the estrogen binding kinetics and stability of the hormone-bound, transcriptionally active "closed" pocket conformation.

On the other hand, White et al. found that the mutation of Tyr541 of the mouse estrogen receptor (Typ537 of the human receptor) to aspartate, glutamate, and alanine increases transcriptional activity without ligand.[31] They proposed that the tyrosine is required to maintain the receptor in a transcriptionally inactive state in the absence of hormone.

The tyrosine is conserved among estrogen receptors but not among steroid receptors. As the result of alignment (Color Plate 2), the residues around Glu542 are conserved among steroid receptors rather than Tyr537. The doublet of Leu/Val-Tyr was observed in the estrogen receptors; however, Phe-Pro occurs in other steroid receptors. It is interesting that the exact sequential position is different, but phenylalanine, reported as a hormonal control by White et al.,[31] occupies a similar position in all receptors. Therefore, Tyr537 could be important for the position or conformational change of C-terminal helix 12, and Glu542 may be playing a role with Tyr537.

Residues Involved in Dimerization

To assess the possible involvement of a residue in dimerization, differences between the SA of residues of a monomeric and of a dimeric form were calculated (Table 4, Color Plate 6). Because the environment of residues whose SA is less than 0.10 was defined as likely to be similar in the monomeric and dimeric forms, Table 4 shows the residues whose SA are $\geqq$0.10 and their 3D situation is indicated in Color Plate 6. The portion of the steroid receptors that is involved in dimerization is similar to that shown in Color Plates 4a and 4e, in which few highly conserved residues were located. The most conserved residues are not involved in dimerization. It was previously reported that the 510th to 522nd residues of mouse estrogen receptor (corresponding to residues 506 to 518

Table 4. Differences in the solvent accessibility of monomeric and dimeric receptors

Estrogen receptor (complex with 17 β-estradiol)							
		1A52			1ERE		
Number	Residue	SA[a] of dimer	SA[a] of monomer	ΔSA[b]	SA[a] of dimer	SA[a] of monomer	ΔSA[b]
381	Cys	0.31	0.41	0.10			
427	Met	0.18	0.34	0.16	0.28	0.51	0.23
430	Ala	0.17	0.56	0.39	0.27	0.61	0.34
431	Thr				0	0.10	0.10
433	Ser	0.39	0.49	0.10			
434	Arg	0.31	0.99	0.68	0.17	0.87	0.70
437	Met				1.12	1.33	0.21
455	Asn	0.01	0.35	0.34	0.03	0.45	0.42
459	Tyr[c]	0.13	1.50	1.37	0.25	1.64	1.39
460	Thr[c]	0.42	0.82	0.40	0.68	0.83	0.15
472	Lys	0.38	0.40	0.20	0.56	0.66	0.10
476	His	0.26	0.75	0.49	0.24	0.75	0.51
479	Leu	0.05	0.15	0.10	0.06	0.18	0.12
480	Asp	0.27	1.02	0.75	0.28	0.98	0.70
483	Thr	0.04	0.41	0.37	0.03	0.38	0.35
484	Asp	0.22	0.85	0.63	0.30	0.94	0.64
487	Ile	0.11	0.40	0.29	0.16	0.45	0.29
497	Leu	0.80	1.11	0.31	0.85	1.30	0.45
498	Gln	0.97	1.29	0.32	0.91	1.10	0.19
501	His	0.12	1.15	1.03	0.29	1.08	0.79
502	Glu	0.32	1.07	0.75	0.33	0.88	0.55
504	Leu	0	0.24	0.24	0	0.30	0.30
505	Ala	0.02	0.48	0.46	0.02	0.50	0.48
506	Gln	0.39	0.76	0.37	0.43	0.77	0.34
508	Leu	0.04	0.62	0.68	0.04	0.60	0.56
509	Leu	0.07	1.30	1.23	0.09	1.21	1.12
510	Ile	0.01	0.14	0.13			
511	Leu[d]	0.03	0.20	0.17	0.02	0.19	0.17
512	Ser	0.03	0.81	0.78	0.04	0.76	0.72
513	His	0	0.73	0.73	0.09	1.16	1.17
515	Arg	0.02	0.71	0.69	0.09	0.87	0.78
516	His	0.16	0.89	0.73	0.32	0.94	0.62
519	Asn	0.12	0.83	0.71	0.10	0.79	0.69
520	Lys				0.54	0.97	0.43
523	Glu	0.80	1.31	0.51	0.54	1.01	0.47
547	His				0.36	0.60	0.24
548	Arg				1.01	1.12	0.11

Progesterone receptor (complex with progesterone)				
		1A28		
Number	Residue	SA[a]of dimer	SA[a] of monomer	ΔSA[b]
885	Lys	0.68	1.11	0.43
886	Gln	0.76	0.93	0.17
889	Leu	0.22	0.89	0.67
892	Leu	0	0.27	0.27
893	Asn	0.38	0.76	0.38
895	Phe	0.35	0.51	0.16
896	Ile	0.31	1.23	0.92
897	Gln	0.69	0.90	0.21
899	Arg	2.02	2.20	0.18
914	Ala	0.17	0.65	0.48
918	Pro	0	0.67	0.67
919	Lys	0.76	0.95	0.19
921	Leu	0.03	0.32	0.29
922	Ala	0.03	0.71	0.68
924	Met	0.77	1.02	0.25

[a] Solvent accessibility; the number of water molecules in contact with the residues.
[b] ΔSA is calculated as the difference between the residues' SA of monomer and dimer.
[c] Residue contained in the variable loops.
[d] Highly conserved residue.

of the human estrogen receptor) were essential for dimerization.[22,32] According to the report,[32] single amino acid substitutions prevented receptor dimerization in the N-terminal half but not in the C-terminal half of the sequence they targeted (residues 505 to 539 of the mouse estrogen receptor, corresponding to residues 501 to 535 of the human estrogen receptor and to residues 867 to 901 of progesterone receptor). Many of the residues of 1A52 potentially involved in dimerization in this study coincide with the conclusion of Fawell et al. However, in the case of 1A28, many of the residues for dimerization are located in the C-terminal half of the targeted sequence of Fawell et al.[22,32]

In light of the associated large ΔSA, the ligand-binding domain of the estrogen receptor (1A52) likely has Tyr459, His501, and Leu509 at its center (shown as red in Color Plates 6c and e), whereas the relatively low ΔSA residues support dimerization (Table 4 and Color Plate 6). This description is applicable to 1ERE with small variation. The three residues are most important, and the location of the C-terminal helix is not involved in dimerization. However, no residues of the ligand-binding domain of the progesterone receptor (1A28) have ΔSA greater than 1.00. As illustrated in Color Plate 5, the progesterone receptor and the estrogen receptor use particular sites for dimerization, and the difference between these receptors probably involves variable residues. Because of the differences in the dimerization sites, the types of binding may differ between the estrogen and progesterone receptors. In the dimerization sites, Leu511 of 1A52, which is one of the highly conserved residues, was potentially involved, but the role of this residue must not be related to contact because Leu511 lies far from the receptor surface (Color Plate 6e). Interestingly, Tyr459, a component of the internal variable loop (Color Plate 2), was a primary interaction residue in the only estrogen receptor. Thus this loop may be important to a receptor-specific function.

Role of Conserved Core Structure and Structural Motif

As described in the preceding sections, the highly conserved residues of the ligand-binding domains of steroid receptors form a core structure. Most of them are buried but some residues are exposured on the surface. They were distributed around the ligand-binding pocket (Leu349, Leu391, Arg394, Phe404, Trp755 and Trp383 of the human estrogen receptor), near the HSP90-binding residues (Leu408, Glu444, Lys449, Leu453, Leu454 and Leu469 of the human estrogen receptor), and in the autonomous transcriptional activation domain (Leu319, Glu323 and Pro324 of it) and between the sites. This may indicate that conserved residues among the steroid receptors are important for their conformational change on binding to or releasing from ligand, HSP90 or the coactivators.

As shown in the previous section, conserved residues are widely distributed in the protein sequence, but make a contiguous 3D structure. Because it is probably important for their biochemical function, the core structure may comprise a structural motif of steroid receptors. PROSITE, the protein sequential motif database, can be used to predict protein function, whereas the method to detect local sequential homology has limited application. The most limiting factor is the numbers of motifs already reported. If conserved residues often make a core structure, the location of residues can be applied to develop a new predictable motif for the function of a protein that has known 3D structure.

CONCLUSION

Analysis of the ligand-binding domains of steroid receptors revealed that most of the highly conserved residues create a core structure. Some of them form a site that recognizes a common structural feature of the steroidal ligands. Several highly conserved residues are located on one side of the surface of the receptor, and they have little involvement in dimerization. Although residues conserved among proteins often are directly attributed with functional roles, assessing the location of these residues in the 3D structure would yield more pertinent functional information, especially among evolutionarily divergent proteins. The structural location of highly conserved residues may relate to the structural change of steroid receptors from inactive to active form. From an another point of view, the core structures can be considered to be “structural motifs”, defined as three dimensionally contiguous structures. In addition, to estimate which residues are involved in dimerization, a quantitative analysis using solvent accessibility was proposed. It can be useful for generally choosing candidates involved in not only in dimerization but also in formation of complexes with other proteins.

ACKNOWLEDGMENT

I acknowledge the MAFF DNA Bank for the use of the GCG program package and Dr. C. Sander of European Bioinformatics Institute who permitted the user license for DSSP. Special thanks go to Dr. Nobuhiro Minaka of National Institute of Agro-environmental Sciences for primary correction of this manuscript. And thanks go to the editor and a referee of this manuscript for their critical reading and many important opinions.

REFERENCES

1 Lubahn, D.B., Joseph, D.R., Sar, M., Tan, J., Higgs, H.N., Larson, R.E., French, F.S., and Wilson, E.M. The rat androgen receptor: primary structure, autoregulation of its messenger ribonucleic acid, and immunocytochemical localization of the receptor protein. *Mol. Endocrinol.* 1988, **2**, 1265–1275

2 Greene, G.L., Gilna, P., Waterfield, M., Baker, A., Hort, Y., and Shine, J. Sequence and expression of human estrogen receptor complementary DNA. *Science* 1986, **231**, 1150–1154

3 Kastner, P., Krust, A., Turcotte, B., Stropp, U., Tora, L., Gronemeyer, H., and Chambon, P. Two distinct estrogen-regulated promoters generate transcripts encoding the two functionally different human progesterone receptor forms A and B. *EMBO J.* 1990, **9**, 1603–1614

4 Arriza, J.L., Weinberger, C., Cerelli, G., Glaser, T.M., Handelin, B.L., Housman, D.E., and Evans, R.M. Cloning of human mineralocorticoid receptor complementary DNA: structural and functional kinship with the glucocorticoid receptor. *Science* 1987, **237**, 268–275

5 Hollenberg, S.M., Weinberger, C., Ong, E.S., Cerelli, G., Oro, A., Lebo, R., Thompson, E.B., Rosenfeld, M.G., and Evans, R.M. Primary structure and expression of a functional human glucocorticoid receptor cDNA. *Nature* 1985, **318**, 635–641

6 Escriva, H., Safi, R., Hanni, C., Lamglois, M.-C., Saumitou-Laprade, P., Steheline, D., Capron, A., Pierce,

R., and Laudet, V. Ligand binding was acquired during evolution of nuclear receptors. *Proc. Natl. Acad. Sci. USA* 1997, **94**, 6803–6808
7 Baker, M.E. Steroid receptor phylogeny and vertebrate origins. *Mol. Cell Endocrinol.* 1997, **135**, 101–107
8 Parker M.G., Steroid and related receptors. *Curr. Opin. Cell Biol.* 1993, **5,** 499–504
9 Bairoch, A., and Bucher, P. PROSITE: Recent developments *Nucleic Acids Res.* 1994, **13**, 3583–3589
10 Tanenbaum, D.M., Wang, Y., Williams, S.P., and Sigler, P.B. Crystallographic comparison of the estrogen and progesterone receptor's ligand-binding domains. *Proc. Natl. Acad. Sci. USA* 1998, **95**, 5998–6003
11 Williams, S.P., and Sigler, P.B. Atomic structure of progesterone complexed with its receptor. *Nature* 1998, **393**, 392–396
12 Cadepond, F., Schwerzer-Groyer, G., Segard-Maurel, I., Jibard, N., Hollenberg, S.M., Giguere, V., Evans, R.M., and Baulieu, E.-E. Heat shock protein 90 as a critical factor in maintaining glucocorticosteroid receptor in a nonfunctional state. *J. Biol. Chem.* 1991, **266**, 5834–5841
13 Xu, M., Dittmar, K.D., Giannoukos, G., Pratt, W.B., and Simons, S.S., Jr. Binding of HSP90 to the glucocorticoid receptor requires a specific 7-amino acid sequence at the amino terminus of the hormone-binding domain. *J. Biol. Chem.* 1998, **273**, 13918–13924
14 Smithies, O., Engels, W.R., Devereux, J.R., Slinghtom, J.L., and Shen, S. Base substitutions, length differences and DNA strand asymmetries in the human G-gamma and A-gamma fetal globin gene region. *Cell* 1981, **26**, 345–353
15 http://www.gcg.com/, Genetic Computer Group, 575 Science Drive Madison, Wisconsin 53711 USA
16 Kabsch, W., and Sander, C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 1983, **22**, 2577–2637
17 Brzozowski, A.M., Pike, A.C., Dauter, Z., Hubbard, R.E., Bonn, T., Engstrom, O., Ohman, L., Greene, G.L., Gustafsson, J.A., Carlquist, M. Molecular basis of agonism and antagonism in the oestrogen receptor. *Nature* 1997 **389**, 753–758
18 Pike, A. C. W., Brzozowski, A. M., Hubbard, R. E., Bonn, T., Thorsell, A.-G., Engstrom, O., Ljunggren, J., Gustaffson, J.-A., and Carlquist, M.: Structure of the ligand-binding domain of oestrogen receptor beta in the presence of a partial agonist and a full antagonist *EMBO J.* 1999, **18,** 4608–4618
19 Shiau, A. K., Barstad, D., Loria, P. M., Cheng, L., Kushner, P. J., Agard, D. A., and Greene, G. L.: The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen. *Cell* 1998, **95,** 927–937
20 Mueller-Fahnow, A., and Egner, U. Ligand-binding domain of estrogen receptors. *Curr. Opin. Biotech.* 1999, **10**, 550–556
21 Danielian, P.S. White, R. Lees, J.A., and Parker, G. Identification of a conserved region required for hormone dependent transcriptional activation by steroid hormone receptors. *EMBO J.* 1992, **11**, 1025–1033
22 Lees, J.A., Fawell, S.E., White, R., and Parker, M.G. A 22-amino-acid peptide restores DNA-binding activity to dimerization-defective mutants of the estrogen receptor. *Mol. Cell. Biol.* 1990, **10**, 5529–5531
23 Tora, L., White, J., Brou, C., Tasset, D., Webster, N., Scheer, E., and Chambon, P. The human estrogen receptor has two independent non-acidic transcriptional activation functions. *Cell* 1989, **59**, 477–487
24 Pierrat, B., Heery, D.M., Chambon, P., and Losson, R. A highly conserved region in the hormone-binding domain of the human estrogen receptor functions as an efficient transactivation domain in yeast. *Gene* 1994, **143**, 193–200
25 Milhon, J., Kohli, K., and Stallcup, M.R. Genetic analysis of the N-terminal end of the glucocorticoid receptor hormone binding domain. *J. Steroid Biochem. Mol. Biol.* 1994, **51**, 11–9
26 Norris, J.D., Fan, D., Kerner, S.A., and McDonnell, D.P. Identification of a third autonomous activation domain within the human estrogen receptor. *Mol. Endocrinol.* 1997, **11**, 747–54
27 Mak, H.Y., Hoare, S., Henttu, P.M.A., and Parker, M.G. Molecular Determinants of the estrogen receptor-coactivator interface. *Mol. Cell. Biol.* 1999, **19**, 3895–3903
28 Beato, M., Herrlich, P., and Schutz, G. Steroid-hormone receptors—many actors in search of a plot. *Cell* 1995, **83**, 851–857
29 Glass, C., Rose, D.W., and Rosenfeld, M.G. Nuclear receptor coactivators. *Curr. Opin. Cell Biol.* 1997, **9**, 222–232
30 Yudt, M.R., Vorojeikina, D., Zhong, L., Skafar, D.F., Sasson, S, Gasiewicz, T.A., and Notides, A.C. Function for estrogen receptor tyrosine 537 in hormone binding, DNA binding, and transactivation. *Biochemistry* 1999, **38**, 14146–14156
31 White, R., Sjoberg, M., Kalkhoven, E., and Parker, M.G. Ligand-independent activation of the oestrogen receptor by mutation of a conserved tyrosine. *EMBO J.* 1997, **16**, 1427–1435
32 Fawell, S.E., Lees, J.A., White, R., and Parker, M.G. Characterization and colocalization of steroid binding and dimerization activities in the mouse estrogen receptor. *Cell* 1990, **60**, 953–962