Molecular and Cellular Endocrinology 334 (2011) 31–38

Contents lists available at ScienceDirect

Molecular and Cellular Endocrinology

journal homepage: www.elsevier.com/locate/mce

Review

Evolution of steroid receptors from an estrogen-sensitive ancestral receptor

Geeta N. Eick[a], Joseph W. Thornton[a,b,*]

[a] Center for Ecology and Evolutionary Biology, University of Oregon, Eugene, OR 97403, USA
[b] Howard Hughes Medical Institute, University of Oregon, Eugene, OR 97403, USA

ARTICLE INFO

Article history:
Received 26 March 2010
Received in revised form 24 June 2010
Accepted 6 September 2010

Keywords:
Evolution
Steroid receptors
Estrogen receptor
Steroid hormones
Testosterone
Progestagens
Androgen receptor
Progesterone receptor

ABSTRACT

Members of the steroid hormone receptor (SR) family activate transcription from different DNA response elements and are regulated by distinct hormonal ligands. Understanding the evolutionary process by which this diversity arose can provide insight into how and why SRs function as they do. Here we review the characteristics of the ancient receptor protein from which the SR family descends by a process of gene duplication and divergence. Several orthogonal lines of evidence – bioinformatic, phylogenetic, and experimental – indicate that this ancient SR had the capacity to activate transcription from DNA estrogen response elements in response to estrogens. Duplication and divergence of the ancestral SR gene subsequently generated new receptors that were activated by other steroid hormones, including progestagens, androgens, and corticosteroids. The androgen and progesterone receptors recruited as their ligands steroids that were previously present as biochemical intermediates in the synthesis of estrogens. This process is an example of molecular exploitation—the evolution of new molecular interactions when an older molecule, which previously had a different function, is co-opted as a binding partner by a newly evolved molecule. The primordial interaction between the ancestral steroid receptor and estrogens may itself have evolved due to an early molecular exploitation event.

Contents

Steroid receptors (SRs) are ligand-activated transcription factors that belong to the diverse nuclear receptor (NR) superfamily of proteins (Robinson-Rechavi et al., 2003). SRs mediate the long-term effects of steroid hormones on reproduction, behavior, immunity, stress responses, and development. Like other NRs, SRs have a highly conserved DNA-binding domain (DBD), which recognizes target DNA sequences (response elements) in the promoter regions of target genes, and a moderately conserved ligand-binding domain (LBD), which mediates hormone binding and ligand-dependent transcriptional activation (Beato et al., 1995). In the absence of ligand, SR LBDs are typically in a transcriptionally inactive conformation. Steroid hormones are fat-soluble derivatives of cholesterol, so they cross-cell membranes, where they bind to their preferred receptor with high affinity and specificity. Upon ligand-binding, the LBD conformation is remodeled and stabilized, forming a new activation surface that can recruit coactivator proteins. SRs also contain a poorly conserved flexible hinge domain and a highly variable N-terminal domain, which contains a second transcriptional activation function (Gronemeyer et al., 2004).

The SR family is a model of protein functional diversification. In most vertebrates, the family includes two receptors for estrogens (ERα and ERβ) and one receptor each for mineralocorticoids (MR), glucocorticoids (GR), androgens (AR), and progestagens (PR). These ligands fall into two major classes: corticosteroids, androgens, and progestagens all have a keto group at the 3-position on

* Corresponding author at: Center for Ecology and Evolutionary Biology, 5289 University of Oregon, USA. Tel.: +1 541 346 0328; fax: +1 541 346 2364.
E-mail address: joet@uoregon.edu (J.W. Thornton).

0303-7207/$ – see front matter © 2010 Published by Elsevier Ireland Ltd.
doi:10.1016/j.mce.2010.09.003

the steroid backbone, while estrogens have a 3-hydroxyl. SRs also regulate largely different sets of target genes by binding to different response elements. ER DBDs bind to an inverted repeat of AGGTCA, whereas those of AR, PR, GR, and MR prefer an inverted repeat of AGAACA; although the DBDs of both kinds of SR can tolerate some variation from their canonical response elements, their specificities are distinct from each other (Zilliacus et al., 1995).

The differences between SR proteins arose by an evolutionary process, so the ultimate reason why SRs have the sequences, structures, and functions that they do lies in their history. In recent years, significant strides have been made in understanding the evolutionary processes by which SRs proliferated and diversified (Laudet, 1997; Thornton, 2001; Thornton et al., 2003; Bridgham et al., 2006; Baker et al., 2007; Baker, 2008; Bridgham et al., 2008; Carroll et al., 2008; Baker and Chang, 2009; Bridgham et al., 2009). Here we review work on the early evolution of the SR family, particularly on the ancestral steroid receptor (AncSR1) from which the entire family is derived. We review knowledge concerning AncSR1and its implications for subsequent SR evolution. We focus specifically on the finding, supported by several lines of evidence, that AncSR1 had intrinsic functions similar to those of modern-day vertebrate estrogen receptors.

1. Origin of the SR family

All members of the SRs descend from a single ancestral receptor, which branched off from the rest of the NR superfamily early in animal evolution (Laudet, 1997; Thornton et al., 2003). The family subsequently proliferated through a series of gene duplications. The phylogeny of the SR gene family (Fig. 1A) indicates that there are two major SR subgroups: ERα and ERβ in one clade, and AR, PR, GR, and MR in the other. Within the latter group, AR/PR and GR/MR are pairs of closely related sister receptors. SR phylogenies with the same general topology as Fig. 1A have been inferred by several other research groups (Baker, 1997; Thornton, 2001; Thornton et al., 2003; Bridgham et al., 2006; Paris et al., 2008; Keay and Thornton, 2009).

Steroid receptors were originally thought to be a vertebrate-specific gene family because of their absence from the genomes of arthropods, nematodes, and urochordates, the first invertebrates to be fully sequenced (Maglich et al., 2001; Dehal et al., 2002; Yagi et al., 2003; Howard-Ashby et al., 2006). That view changed, however, when genes with clear sequence homology to SRs were discovered in mollusks (Thornton et al., 2003; Kajiwara et al., 2006; Keay et al., 2006; Castro et al., 2007; Matsumoto et al., 2007). There are two major clades of bilaterally symmetric animals: the protostomes (including arthropods, nematodes, mollusks, annelids, and many other phyla) and the deuterostomes (which include vertebrates and other chordates, as well as echinoderms). The discovery of SRs in mollusks indicated that the family is in fact as old as the protostome–deuterostome ancestor. More recently, SRs have been isolated from another protostome phylum, the annelids (Keay and Thornton, 2009). All these protostome receptors are most similar in sequence to the vertebrate ERs, with considerably lower similarity to the other SRs and the closest NR outgroup, the ERRs. Phylogenetically, the protostome receptors group with strong support as members of the SR family (Keay et al., 2006; Keay and Thornton, 2009) (Fig. 1A).

SRs proliferated only in the chordates (Fig. 1B). The single ER-like gene appears to be the only SR in protostomes, because no others have been detected in the fully sequenced genomes of mollusks and annelids. Within the chordates, however, the most basally branching subphylum – the invertebrate cephalochordates, also known as lancelets or amphioxus – possess two SRs, one ortholog for each of the two major SR clades, the ERs and the AR/PR/GR/MR group (Bridgham et al., 2008; Katsu et al., 2010). The former has high sequence similarity to the human ERs, whereas the other has approximately equal similarity to the ERs and the other SRs. This result indicates that the initial duplication of AncSR1 to produce the two major SR classes occurred before the last common ancestor of all chordates (Fig. 1A and B). A second duplication – of the ancient gene AncSR2, the last common ancestor of AR, PR, GR, and MR – occurred after vertebrates diverged from cephalochordates. Thus jawless fishes, the earliest branching vertebrates, contain three SRs – one ER, one AR/PR ortholog, and one GR/MR ortholog. Information on conserved syntenic relationships with other gene families (Thornton, 2001) suggests that this duplication may have occurred in the first of two whole-genome duplication events early in the vertebrate lineage (Van de Peer et al., 2009). The six receptors found in most jawed vertebrates today were produced in a final round of duplication, possibly due to the second whole-genome duplication during early vertebrate evolution (Thornton, 2001; Carroll et al., 2008). The absence of SRs from ecdysozoans (arthropods and nematodes (Maglich et al., 2001), echinoderms (Howard-Ashby et al., 2006), and the urochordates (Dehal et al., 2002)) indicates that the progenitor SR gene was lost independently during the evolution of these taxa.

The gene family phylogeny on which these inferences are made is well supported at most nodes (Fig. 1A). The major exception is the position of the protostome SRs as a sister group to the vertebrate ERs, which is weakly supported. The maximum likelihood topology implies that the initial duplication of AncSR1 occurred before the protostome–deuterostome divergence, and the ortholog of the AR/PR/GR/MR clade was then lost in the protostomes (Fig. 1B, scenario *a*). An alternate arrangement is also possible, however, in which the protostome receptors are equally orthologous to the entire SR family, having branched from the rest of the family before the duplication of AncSR1 to yield the family's two major clades (Fig. 1). This scenario (*b* in Fig. 1B) implies that the initial duplication of the ancestral steroid receptor to produce the two major clades of SRs occurred in the deuterostomes, after their divergence from protostomes. In both cases, AncSR1 is the last common ancestral gene from which all extant SRs descend. In scenario *a*, AncSR1 represents the single SR gene at the time of its duplication into the ER and AR/PR/GR/MR lineages, before the last common ancestor of protostomes and deuterostomes. In scenario *b*, AncSR1 represents the single SR gene in the protostome–deuterostome ancestor.

2. Functions of the ancestral steroid receptor: evidence from extant taxa

The similarity of all the invertebrate receptor sequences to the vertebrate ERs suggests that the ancestral SR protein was likely to have been ER-like in both sequence and in function. Analyses of the functions of SR proteins in a phylogenetic context corroborate this view. They have also revealed several SR family members with unusually divergent functions.

The capacity to bind estrogens and estrogen response elements is clearly as old as the common ancestor of protostomes and deuterostomes. Like those of vertebrate ERs, the DBDs of all annelid, mollusk, and cephalochordate steroid receptors all bind to and regulate transcription from EREs (Keay et al., 2006; Bridgham et al., 2008; Keay and Thornton, 2009; Katsu et al., 2010). The hypothesis that AncSR1 bound EREs would explain the DNA-binding behavior of the entire SR family as due to direct descent from the ancestor, with a single shift in the AR/PR/GR/MR lineage from preferring AGGTCA binding sites to AGAACA.

As for the ligand-binding domain, most of the invertebrate SR family members are also ER-like, but some have unique functionalities. For example, steroid receptors from two annelid species have