Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochemical and Biophysical Research Communications 306 (2003) 1–4
BBRC
www.elsevier.com/locate/ybbrc

Breakthroughs and Views

DNA binding of nuclear hormone receptors influences their structure and function

E. Brad Thompson* and Raj Kumar

The University of Texas Medical Branch, Department of Human Biological Chemistry and Genetics, Galveston, TX 77555-0645, USA

Received 25 April 2003

The nuclear hormone receptors (NHRs) comprise a large family of proteins that regulate gene transcription. Many NHRs are activated to do this by the binding of receptor type-specific ligands, such as the steroid hormones and vitamin D, thyroid hormone, retinoids, bile acids, and certain lipids. Other NHRs, known as orphan receptors, have no known ligand. The NHRs act on specific genes by binding as homo- or hetero-dimers to regulatory element (RE) sites in the DNA, usually found at varying distances upstream from the relevant transcription start site. The classic specific REs have been extensively characterized. They contain a core sequence of five or six nucleotides with a characteristic consensus sequence for each receptor type. Receptor-specific core elements are arranged in the classic REs in tandem arrays, with each core sequence (half-site) in the array separated by 0–6 non-specific nucleotides and aligned variously, so that their paired sequences read head-to-head, head-to-tail, etc. The core sequence and topography are characteristic for the specific receptor type, with some overlaps [1–3]. However, these ''rules,'' particularly those of the DNA sequence in the cores, are sometimes not met by the REs found in actual gene regulatory regions. The nucleotides found in the core DNA sequences may deviate significantly from the consensus sequence [4,5]. Consensus sequences, after all, are compilations of the most frequently used bases at each position, not necessarily the sequence found in a particular natural gene. Also, it should be recalled, the DNA sequences flanking the RE and the location of the RE with respect to the gene it regulates are important for the overall affinity of the NHR for its RE [5–10].

Other ways that NHRs find their way to bind at the regulatory regions of specific genes involve interactions with other, heterologous transcription factor protein sites that include a partial, half-site RE for the NHR, alongside a binding site for the heterologous factor, or at sites at which the NHRs are exclusively bound by way of the heterologous factor at its own cognate DNA site [5,6,11,12].

NHRs can be seen to comprise a family due to their modular structure. This consists of well-recognized and very closely related DNA binding domains (DBDs), less closely related but recognizable ligand binding domains (LBDs) that invariably form the C-terminal part of the protein, and quite diverse amino terminal domains (NTDs). NTDs vary in length from a relatively short stretch preceding the DBD to hundreds of amino acids [1]. The longer NTDs as found in the steroid hormone receptors contain powerful transcription transactivating domains. As yet no NTD structures have been ascertained, because when expressed as recombinant proteins, these domains fail to demonstrate organized structures. On the other hand, recombinant DBDs and LBDs express as globular proteins, and this has allowed resolution of their structures by NMR and/or crystallography. These structures show that among the DBD and LBD domains, individual examples follow a clear structural pattern [13,14].

The classic model for NHR action holds that NHRs are bound to DNA as homo- or hetero-dimers at specific REs by virtue of high affinity for the relatively short sequences contained therein. Once bound, the receptor collects a variety of ancillary factors which modify chromatin and/or contact the multiprotein primary transcription complex, so as to repress or enhance transcription from the relevant promoter. Over recent

* Corresponding author. Fax: 1-409-772-5159.
E-mail address: bthompso@utmb.edu (E.B. Thompson).

doi:10.1016/S0006-291X(03)00877-5

years, immense effort has led to the identification of an ever increasing number of protein and RNA ancillary factors [15,16]. Many of these factors have been shown to interact directly with the NHRs. Therein lies a problem, since it is not clear how the available surface area on an NHR can accommodate specific sites for all the factors. Often this difficulty is dealt with by noting that many such factors may be expressed in a "cell specific manner." Yet even if only the ubiquitously or commonly expressed factors are considered, they are sufficient in number to raise the hypothetical problem. And even if each type of NHR were to find only a unique cell-specific set of ancillary factors, there still would be a problem, since the fully folded NHR would have to be able to bind a variety of heterologous proteins which do not contain the same NHR-binding motifs.

Models that attempt to explain the specificity of action and interaction between NHRs and ancillary proteins, and consequently of the ability of NHRs to regulate specific genes, have another interesting set of facts to deal with. The REs in actual genes often vary from the exact consensus sequence. In fact, in many actual genes the RE for a specific receptor may deviate so far from the consensus sequence that it cannot be recognized by computer software designed to find such sequences. In some such sites, it has been shown that ancillary proteins are absolutely required for the binding of the NHR [5,6]. In others, the site simply does not exactly have the consensus sequence and yet functions in vivo to bind its NHR so as to provide gene regulation. The molecular mechanisms that explain this have not been fully defined.

In sum, the NHRs contain three regions: the NTD (variable sequences and length), the DBD (most conserved), and the LBD (a 12-helix globular domain, somewhat variable in sequence, slightly in length). Each of these has been shown to be capable of interactions with other proteins. Binding of ligands to the LBD causes a rearrangement of its C-terminal helix (usually helix 12) so as to form a surface to which co-factor proteins bind [13]. These may modify chromatin and also provide platforms for additions to the multi-protein complexes [15,16]. Ultimately such complexes interact with the TFIID and RNA polymerase multi-protein complexes. When expressed as recombinant proteins, however, the longer NTDs found in some NHRs show no specific structure, and as such interact weakly at best with co-factor proteins. But when these NTDs are caused to fold, they show ability to bind such factors [17–22]. It is unknown how these NTD-based multi-protein complexes actually assemble, how the various heterologous co-factor proteins fit on the NHR surface, exactly how chromatin is "remodeled" by them, or how the primary transcription complexes are contacted, physically.

The nucleotide sequence in an RE may affect NHR structure

One hypothesis that helps to reconcile these issues states that the exact nucleotide sequence of the RE affects not only the overall affinity of the receptor for its RE site, but also influences the three dimensional configuration of the receptor. In fact, this hypothesis has been suggested to apply to transcriptional regulatory proteins in general [22]. Consequently, in an RE-specific way, the surfaces on the NHR are modified so that various critical ancillary factors can bind. This hypothesis clearly differs from the classic textbook description of how NHRs work. Site-specific DNA binding may be more than the textbook mechanism for localizing the receptor to the right place so as to regulate transcription of the relevant gene. According to the hypothesis, it can also be a trigger for an active intramolecular event that changes the shape of the NHR in a site-specific way. This could result in binding certain ancillary proteins in a site-specific way. Since transcriptional regulation for a specific gene depends upon the interactions of these proteins, the exact DNA sequences of the available RE sites in the regulatory region of the DNA of the gene could help determine gene regulation. In addition to the DNA sequence within the RE, the location of the RE with respect to DNA sites for other factors is often of high importance, since it can determine the availability of other important transcription factors with which the NHR interacts. The combination of all these events may lead to transcription regulation in a cell and promoter specific manner.

It is well known that protein:DNA interactions can change the shape of both the DNA and the bound protein. In specific cases, the changes may be quite profound. For example, when bound at its TATA sequence, the TATA box binding protein (TBP) produces a very sharp bend in DNA. It was reasoned some time ago on thermodynamic principles that DNA binding was very likely to affect the structures of many DNA binding proteins, and as a subset of these, transcription factors [23]. At the time of those calculations little was known about the structures of NHRs except for their DNA binding domains, but it was calculated that even this relatively well-folded domain could be influenced in structure by the act of binding DNA. For NHRs, as several DBD structures were obtained and compared when free or DNA-bound, it was shown that in fact, the predictions were true. Structural changes both at the DNA binding surface and away from it occur upon DBD:DNA binding [24,25]. DNA binding also influences receptor functions further away in the NHR. For the thyroid hormone receptor/retinoid X receptor heterodimer (TR/RXR), it has been shown that varying the sequences in the TRE to which the heterodimer binds modulates the conformation of the heterodimer, with

consequences for the transcriptional activity of the dimer. Specific mutations in the TRE permitted binding but altered transcriptional activity [26]. This result showed both that site-specific DNA binding alone is not sufficient for transcriptional activation by an NHR and that DNA specific sequences in TREs affect the conformation of the heterodimer as revealed by partial proteolysis experiments. The exact whereabouts within the NHR of the regions modified by the DNA binding were not fully ascertained.

In similar studies, specific sequences within estrogen response elements (EREs) have been shown to modulate the conformation of both the estrogen receptor β and α forms [27–29]. Furthermore, these ERE sequence-specific alterations in conformation of the ERE-bound ER homodimer resulted in altered recruitment of co-activator proteins to the ER:ERE. This provides an intellectually satisfying rationale for the varied transcription effects of specific sequences found in the REs controlling specific genes. One can image that, as hypothesized, specific sequences allow binding that changes the surfaces of the ER such that particular co-activator (or co-repressor) proteins are offered binding surfaces on the ER. In the same cell, therefore, availability of proper surfaces on the NHR and the affinities between protein surfaces would determine the selection of the final co-activators found bound in the ER:ERE complex.

Recently, progress has been made as to the NHR regions in which effects of DNA binding on receptor structure occur. Dramatic folding effects have been seen on the NTD when DNA binding takes place. In the subfamily of steroid–hormone receptors in particular, the NTDs, though heterogeneous in primary sequence, all contain one or more powerful transcription transactivation domains. These have been defined on the basis of mutations tested in transfection systems. Whereas the DBD and LBD domains of the NHRs are globular proteins in solution when expressed as recombinant proteins, the NTDs or their transactivation domains show little or only limited structure. Thus they appear to be composed of a large collection of conformers, among which may be a very small proportion with the correct structure for carrying out transcriptional transactivation [30]. The NTD activation domains of the NHRs may be among the ''natively unfolded'' regions which are frequently found in proteins that interact with other proteins. Recent work in which the NTD and DBD of certain receptors have been expressed together as a single large peptide has shown that DNA binding can promote folding in the NTD activation region.

Structural studies with the glucocorticoid receptor showed that stoichiometric binding to a consensus GRE of a two domain fragment of glucocorticoid receptor containing all the natural amino acids from the N-terminus through the DBD led to formation of tertiary structure in the NTD [31]. This result suggests that considerable RE binding energy is devoted to intramolecular rearrangement in the NTD. When the GR was bound to DNA, the NTD became capable of binding known coactivator proteins, as well as TBP. Similar studies on the progesterone receptor indicated that while the NTD contained some structured elements, DNA binding resulted in additional structure in the NTDs of both the A and B forms of the receptor [21,22].

Together, these results suggest that one of the reasons why sequence specific DNA binding has such a profound effect on function of the nuclear hormone receptors in general and the steroid–hormone receptors in particular may be because the relatively unstructured NTD gains structure such that it can bind appropriate co-factors when RE:DBD interactions occur. Considering the previously mentioned studies in which functional DNA sequence-specific effects were shown on the conformation of the TR:RXR heterodimer and on the estrogen receptor, and consequently on the ability of these receptors to acquire co-factors, an attractive

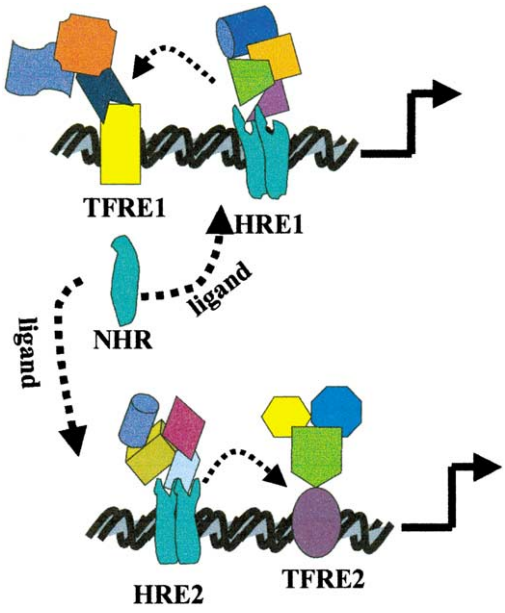

Fig. 1. Model of the influence of DNA sequence on NHR structure and function. An NHR is shown in three forms: unoccupied by ligand, and occupied by ligand and bound as a dimer to two canonical response elements (HRE 1, 2) with differing nucleotide sequence. The bound NHR is shown as a homodimer of differing shapes, dictated by the sequence-specific binding site. Each HRE is shown near an RE for an unrelated transcription factor (TF). The structure of the NHR, influenced by its HRE, results in altered interactions with various co-activators (and repressors) and the constellation of factors formed at the nearby TF RE. These co-factor proteins are shown in different shapes and colors. We propose that the structure of the NHR NTD is especially likely to be influenced by the HRE sequence bound. (For simplicity's sake, the NHRs that form heterodimers on DNA, those that have no ligands or do not require ligands to bind DNA specifically, and the complex of proteins found with unliganded NHRs are not shown. However, the model can extend to all of these).

conclusion is that much of the specificity being imparted by DNA sequences is due to the effects on the fold of the structurally elusive N-terminal activation domain. This is not to say that there are no effects on the structure of the more globular DBD or LBD. As stated above, it is well known that there are effects on the DNA binding domain structure consequent to binding an RE, and it is possible that there will be adjustments in the LBD that result from DNA binding. The effect of DNA sequence, if any, on these effects has yet to be determined in any detail. However, the lesser degree of initial structure that exists in the NTD activation domain, coupled with the changes seen in this domain when DNA binding occurs in at least two examples of the family, suggests that a great deal of the net effect will occur through its altered structure influenced by DNA binding.

We therefore suggest that the textbook models explaining the mechanism of action of the nuclear hormone receptor family need revision to include the fact that the varied RE sequences to which each of this family binds influence the actual structure of the NHR, and especially its NTD. Consequently the ancillary proteins that bind to the NHR are excluded or included in the ultimate complex by virtue of the conformed surfaces available on the NHR consequent to the DNA binding (Fig. 1). Active research is testing this model.

Recent experiments with fluorescent labeled NHRs interacting with chromatin in living cells have suggested that the occupancy time at binding sites on chromatin is relatively short, so that receptors seem to be moving on and off the chromatin DNA in a time frame of seconds [32]. If this is so, and if it is equally so that site- and sequence-specific DNA binding influence the conformation of receptor interactions with ancillary factors/co-factors, it can be seen than an extremely dynamic situation can exist as the receptor functions on chromatin in vivo. Together, the two types of data lead to the idea that NHRs have the capacity to rapidly form and reform multiprotein complexes. This would provide a mechanism for sorting and choosing very specific complexes responsible for gene repression and transactivation.

References

[1] R.M. Evans, Science 240 (1988) 889–895.
[2] D.J. Mangelsdorf, C. Thummel, M. Beato, P. Herrlich, G. Schütz, K. Umesono, B. Blumberg, P. Kastner, M. Mark, P. Chambon, R.M. Evans, Cell 83 (1995) 835–839.
[3] D.J. Mangelsdorf, R.M. Evans, Cell 83 (1995) 841–850.
[4] M. Beato, G. Chalepakis, M. Schauer, E.P. Slater, J. Steroid Biochem. 32 (1989) 737–748.
[5] P.C. Lucas, D.K. Granner, Annu. Rev. Biochem. 61 (1992) 1131–1173.
[6] A.A.J. Scheller, D.M. Robins, Mol. Cell. Biol. 13 (10) (1993) 6326–6335.
[7] P.C. Kulakosky, M.A. McCarty, S.C. Hernigan, K.E. Risinger, C.M. Klinge, J. Mol. Endocrinol. 29 (1) (2002) 137–152.
[8] T.T. Huynh, D.W. Ray, I.J. Brogan, A. Stevens, J.R. Davis, A. White, A. White, J. Endocrinol. 172 (2) (2002) 295–302.
[9] C. Juge-Aubry, A. Pernin, T. Favez, A.G. Burger, W. Wahli, C.A. Meier, B. Desvergne, J. Biol. Chem. 272 (40) (1997) 25252–25259.
[10] J.H. Anolik, C.M. Klinge, C.L. Brolly, R.A. Bambara, R. Hilf, J. Steroid Biochem. Mol. Biol. 59 (5–6) (1996) 413–429.
[11] J.N. Miner, K.R. Yamamoto, Trends Biochem. Sci. 16 (1991) 423–426.
[12] R. Schüle, R.M. Evans, Trends Genet. 7 (1991) 377–381.
[13] R. Kumar, E.B. Thompson, Steroids 64 (5) (1999) 310–319.
[14] M.-J. Tsai, B.W. O'Malley, in: C.C. Richardson, J.M. Abelson, A. Meister, C.T. Walsh (Eds.), Annual Review of Biochemistry, 63, Annual Reviews Inc, Palo Alto, 1994, pp. 451–486.
[15] L. Xu, C.K. Glass, M.G. Rosenfeld, Curr. Opin. Genet. Dev. 9 (2) (1999) 140–147.
[16] J.D. Chen, H. Li, Crit. Rev. Eukaryot. Gene Expr. 8 (2) (1998) 169–190.
[17] R. Kumar, J.C. Lee, D.W. Bolen, E.B. Thompson, J. Biol. Chem. 276 (21) (2001) 18146–18152.
[18] R. Kumar, I.V. Baskakov, G. Srinivasan, D.W. Bolen, J.C. Lee, E.B. Thompson, J. Biol. Chem. 274 (35) (1999) 24737–24741.
[19] J. Reid, S.M. Kelly, K. Watt, N.C. Price, I.J. McEwan, J. Biol. Chem. 277 (22) (2002) 20079–20086.
[20] D.L. Bain, M.A. Franden, J.L. McManaman, G.S. Takimoto, K.B. Horwitz, J. Biol. Chem. 275 (10) (2000) 7313–7320.
[21] D.L. Bain, M.A. Franden, J.L. McManaman, G.S. Takimoto, K.B. Horwitz, J. Biol. Chem. 276 (2001) 23825–23831.
[22] J.A. Lefstin, K.R. Yamamoto, Nature 392 (1998) 885–888.
[23] R.S. Spolar, M.T. Record Jr., Science 263 (1994) 777–784.
[24] P.J. van Tilborg, M. Czisch, F.A. Mulder, G.E. Folkers, A.M. Bonvin, M. Nair, R. Boelens, R. Kaptein, Biochemistry 39 (3) (2000) 8457–8474.
[25] S.M. Holmbeck, H.J. Dyson, P.E. Wright, J. Mol. Biol. 284 (3) (1998) 533–539.
[26] M. Ikeda, E.C. Wilcox, W.W. Chin, J. Biol. Chem. 271 (38) (1996) 23096–23104.
[27] M.A. Loven, J.R. Wood, A.M. Nardulli, Mol. Cell. Endocrinol. 181 (1–2) (2001) 151–163.
[28] M.A. Loven, V.S. Likhite, I. Choi, A.M. Nardulli, J. Biol. Chem. 276 (48) (2001) 45282–45288.
[29] Y.P. Driscoll, J. Huang, S. Bhagat, R. Hilf, R.A. Bambara, M. Muyan, Mol. Endocrinol. 16 (4) (2002) 674–693.
[30] R. Kumar, E.B. Thompson, Mol. Endocrinol. 17 (1) (2003) 1–10.
[31] R. Kumar, I.V. Baskakov, G. Srinivasan, D.W. Bolen, J.C. Lee, E.B. Thompson, J. Biol. Chem. 274 (1999) 24737–24741.
[32] G.L. Hager, Ernst Schering Research Foundation Workshop 40 (2002) 111–129.