AD_____

Award Number: DAMD17-00-1-0048

TITLE: Structure/Function Studies of the Androgen Receptor DNA-Binding Region

PRINCIPAL INVESTIGATOR: Fraydoon Rastinejad, Ph.D.

CONTRACTING ORGANIZATION: University of Virginia
Charlottesville, VA 22904

REPORT DATE: April 2004

TYPE OF REPORT: Final Addendum

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

# REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB No. 074-0188*



| 1. AGENCY USE ONLY (Leave blank) | 2. REPORT DATE | 3. REPORT TYPE AND DATES COVERED |
|---|---|---|
| | April 2004 | Final Addendum (1 Apr 2003 - 31 Mar 2004) |

**4. TITLE AND SUBTITLE**

Structure/Function Studies of the Androgen Receptor DNA-Binding Region

**5. FUNDING NUMBERS**

DAMD17-00-1-0048

**6. AUTHOR(S)**

Fraydoon Rastinejad, Ph.D.

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**

University of Virginia
Charlottesville, VA 22904

E-Mail: fr9c@virginia.edu

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)**

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**

**11. SUPPLEMENTARY NOTES**

Original contains color plates: ALL DTIC reproductions will be in black and white

**12a. DISTRIBUTION / AVAILABILITY STATEMENT**

Approved for Public Release; Distribution Unlimited

**12b. DISTRIBUTION CODE**

**13. ABSTRACT (Maximum 200 Words)**

The androgen receptor (AR) regulates the growth and differentiation of prostate cells and is an important drug target for prostate cancer chemotherapy. The research goals associated with this study are to characterize the structural and functional aspects of the AR in order to uncover the potential of its domains, and in particular the DNA-binding domain, as a drug target. In this final report that includes the addendum period (one year no-cost extension), I discuss the results obtained over the three year course of the funding, plus the additional no cost extension period of one year, towards characterization of the AR and related proteins. Among our findings are a) the DNA-binding domain of the androgen receptor and related nuclear receptors act as their nuclear export signals, b) their export is dependent on their binding to the protein calreticulin in the presence of calcium, c) by analogy with our recent crystal structure of EcR-Usp on DNA, a pair of DNA-binding domains are arranged symmetrically as a homodimer with respect to each other and directly on the half-sites of their target DNA, and the proteins are subject to substantial plasticity in their DNA contacts, and d) the ligand binding domain is analogous to the FXR ligand-binding domain and shares highly related surfaces responsible for DHT binding and coactivator binding.

**14. SUBJECT TERMS**

Prostate Cancer

**15. NUMBER OF PAGES**

97

**16. PRICE CODE**

| 17. SECURITY CLASSIFICATION OF REPORT | 18. SECURITY CLASSIFICATION OF THIS PAGE | 19. SECURITY CLASSIFICATION OF ABSTRACT | 20. LIMITATION OF ABSTRACT |
|---|---|---|---|
| Unclassified | Unclassified | Unclassified | Unlimited |

## Table of Contents

# Structure-Function Studies of Androgen Receptor DNA Binding Region

## INTRODUCTION

The androgen receptor (AR) is a member of the nuclear hormone receptor superfamily, a class of ligand-activated gene-regulatory factors. Nuclear receptors, including AR, regulate gene expression through specific DNA binding sites known as hormone response elements. The nuclear receptors consist of modular domains, including a central and well-conserved region known as the DNA-binding domain (DBD), an un-conserved C-terminal extension known as the hinge region, and a C-terminal portion known as the ligand-binding domain (LBD) [see Figure 1]. The DBD imparts the receptor's entire DNA-binding specificity and selectivity. The LBD allows receptors to bind to ligands and coregulator proteins. The AR is a good drug target for current and future anti-prostate cancer therapy, because it is responsible for the growth and differentiation of the normal and cancerous cells of the prostate.

The DBDs within the nuclear receptor family may be classified into distinct groups based on the geometry of their target DNA binding sites (also known as hormone response elements). Class-I receptors bind to hormone response elements that consist of an inverted (or palindromic) repeat of the hexamer AGGTCA or AGTTCA (see Figure 1). In this group is AR homodimer, the glucocorticoid receptor (GR) homodimer, the mineralocorticoid receptor (MR) homodimer, and the Ecdysone receptor-Ultraspiracle (EcR-Usp) heterodimer. Class-II receptors are those that bind to direct repeats of the same hexameric sequences. These members are generally the receptors that form heterodimers with the retinoid-X-receptor (RXR), such as the thyroid hormone receptor (TR), the vitamin D receptor (VDR), and the all-trans retinoic acid receptor (RAR). Our lab has previously characterized a number of new crystal structures for the type-II receptors in which the DBDs of RXR and its partners were assembled on direct-repeat hormone response elements.

Prior to starting the studies in this proposal, the DBDs had been

viewed as having only one type of function, that of targeting the receptors to their correct DNA sites. Therefore, we sought to investigate other possible functions associated with the DBDs, and uncovered a previously unknown role in nuclear export. We also characterized and reported on the other protein players responsible for the export of AR. These studies have produced significant impact on the field, as there was no previous knowledge available regarding how nuclear receptors exit the nucleus.

In terms of the characterizing the structures of the DBDs, we have been pursuing the goal of providing a detailed crystallographic model of how the AR DBD and in general the other class-I receptors are able to form homodimers on their consensus type response element. Due to intractable diffraction limitations with respect to the AR complex, we instead pushed forward and determined another class-I receptor/DNA complex structure that we know shares all of the essential hallmarks associated with the AR/DNA complex. In addition, we went beyond all of the aims in the original proposal by attempting to extend the known structural features reported by another group on the AR ligand-binding domain (LBD). In that structure, there was no coactivator present in

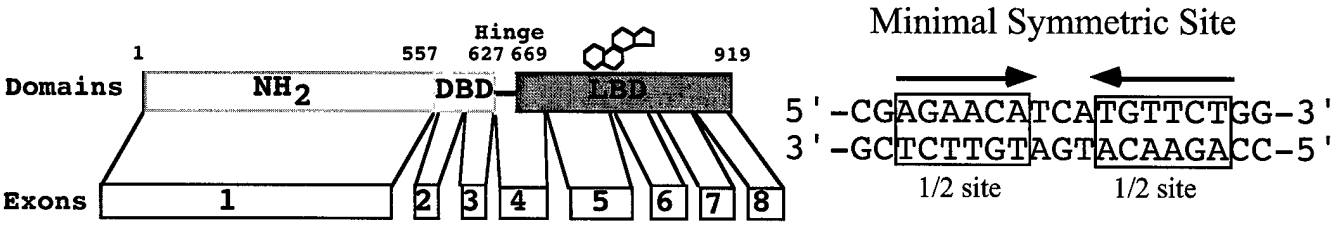



Figure 1: The domain structure of the androgen receptor and the inverted-repeat geometry and sequence of its hormone response element (DNA-binding site).

the crystal structure and thus no lessons as to how a steroidal molecule activates the AR. Again, by focus on a highly related steroid receptor, FXR, we were able to gain important new structural insights as to how steroids are likely to reshape the LBD configuration to recruit the binding of LXXLL motifs associated with the p160 protein of coactivators.

BODY and Key Research Accomplishments

**DNA-Binding Domain Functional Characterization**
(Appendix-1, Black et al., 2001) contains our first reported findings. In this study, we were able to identify the location of the nuclear export signals (NES) within the AR, and also within closely related receptors and to characterize their functions related to their nuclear export pathway. We found that the core DBD region of AR and this receptor, which are closely homologous to each other but unrelated to any known NES, are the regions that are directly responsible for their nuclear export. Mutational analysis revealed that a 15 amino acid sequence between the two zinc binding loops in the DBD confers nuclear export to a green-fluorescence protein construct, and alanine-scanning mutagenesis was used to identify the residues within this sequence that are critical for export. We went on to test and show that the DBDs from ten different nuclear receptors all function as the same export signal. Our findings guided us to propose that NLS-mediated import and DBD-mediated export define a shuttling cycle that integrates the compartmentalization and activity of nuclear receptors.

(Appendix-2, Holaska et al, 2002) contains our second published report which describes how we have recently characterized a pathway for nuclear export of the nuclear receptors in mammalian cells. This pathway involves the Ca(2+) -binding protein calreticulin (CRT), which directly contacts the DNA binding domain (DBD) of AR, and related receptors and allows their transport from the nucleus to cytoplasm. We examined a role of Ca(2+) in CRT-dependent export of these receptors. Removing Ca(2+) from CRT inhibits the stimulation of the nuclear export, an effect due to the failure of CRT to bind the DBD of these receptors. These effects are reversible, when there is a restoration of Ca(2+). Depletion of intracellular Ca(2+) inhibits receptor export in intact cells under conditions that do not inhibit other nuclear transport pathways. We also found that the Ran GTPase is

not required for GR export. It appears that the nuclear export pathway used by steroid hormone receptors are most likely distinct from the Crm1 pathway. Our data has been leading us to consider whether signaling events that increase Ca(2+) could positively regulate CRT and inhibit GR function through nuclear export.

**DNA-Binding Domain Structural Characterization**

We obtained crystals of the AR-DBD/DNA which were small and generally poorly diffracting (see Figure 2). To improve the quality of the crystals and to obtain a high-resolution structure, we attempted a series of different approaches. First, we made a series of new protein expression constructs. The rationale was that perhaps additional sequences C-terminal to the core DBD would enhance DNA binding and stabilize the complex and the crystal lattice. Five different constructs were made in our laboratory, each with a common N-terminal (His)6 tag for purification. Each of the five new constructs contained the core DBD region, plus 0, 10, 20, 30, 41 residues of hinge regions, respectively. The purification involved an initial Ni-NTA column followed by a S-column (cation exchange). These constructs are cloned into pET16b vector, and expressed in BL21DE3 strain. A Ni-NTA column and an S column are used to purify the proteins.

In each case, we did obtain good protein yields, and could purify the AR-DBDs effectively as described above (see Figure 4 for an example). Moreover, we tested and attempted to confirm that each of these AR DBD polypeptides retained the ability to bind DNA (see Figure 5). A surprising finding was that in essentially all cases, the binding to an Androgen Response element (ARE) could be competed out with excess non-specific DNA (Figure 5). This was disturbing and suggested that the reason we may have been unable to get well-

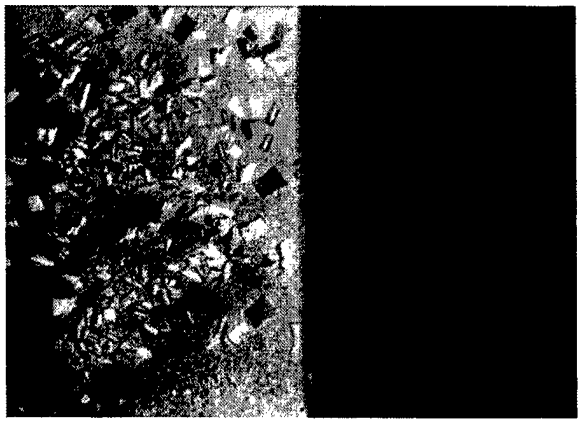


Figure 2: Preliminary crystals of the AR-DBD in complex with a consensus inverted-repeat-3 DNA diffracted poorly when tested at home with a Rigaku-RU2HR X-ray generator B: crystals of the 'extended' complex involving the 27 base pair natural (probasin gene) AR response element, with AR DBD with a longer hinge region also diffracted poorly.

diffracting crystals was the AR-DBD was binding the synthetic DNA in the crystallization mixture using a non-specific binding mode.

We then set out to vary the type of DNA used in the crystallization with the hope that we could identify a more suitable binding site which would promote very tight and specific binding with the AR-DBD. In particular, we examined the probasin gene which contains more than one AR binding element. Using additional DNA synthetic constructs corresponding to these AREs, we again were unable to show very strong and specific binding which was not easily competed out with non-specific DNA.

These efforts to produce stable protein-DNA complexes of AR continued without much success over the first two years. In the past year, we switched our strategy by focusing instead on another steroid receptor DBD /DNA complex which we felt shared most of the hallmarks that are associated with AR. This work has recently led to our determination and refinement of a high resolution (2.4 Å) crystal structure of the EcR-Usp heterodimer bound to a symmetric DNA site related to the Androgen-response elements.

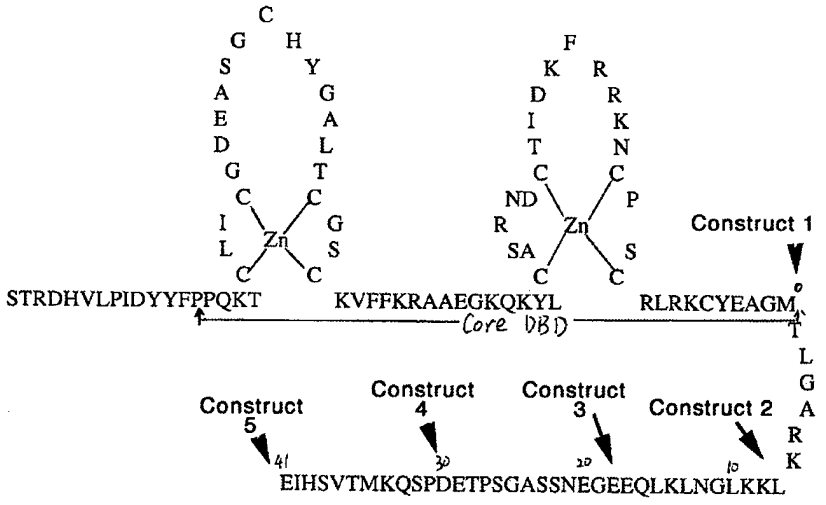



Figure 3: AR-DBD protein constructs made to improve crystal quality and diffraction.

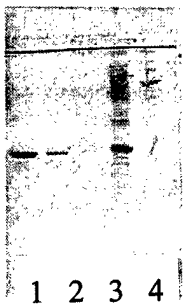



Figure 4: typical protein purification of AR-DBDs. Construct 5 is purified from a Pharmacia Fast-S column Lanes 1-3 show the peak fractions eluted with 300 mM NaCl containing buffer. Lane 4 show molecular weight markers.

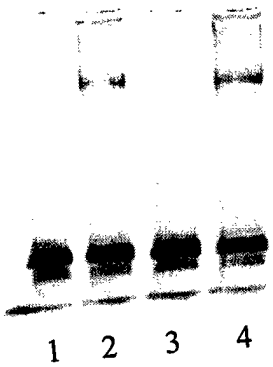



Figure 5: Gel shift analysis of the binding of Construct 5 (see Figures 3 and 4 to DNA. Lane 1: Free DNA (ARE2; 40nM); lane 2: DNA: AR-DB at ratio of 1: 15 (mol ratio) lane 3: Same as lane 2 but with 2uM poly dIdC; lane 4: ARE DNA to AR-DBD ratio of 1: 30.

Figure 6 shows our recently determined crystal structure of the EcR-Usp complex bound to a synthetic symmetric DNA element (see Appendix 3, Devarakonda *et al.*, 2003)– having the same half-site geometry and arrangement as a consensus ARE. The following lessons learned from the analysis of this complex are expected in the AR DBD complex on its symmetric element. First, the DBDs use an alpha-helix to engage the major groove of their core hexameric site. These portion of the receptor DBD, ("the recognition helix') is nearly identical in amino-acid sequence between the AR, Usp and Ecr. Second, the subunits form a dimeric arrangement in which the Zn-II loops of the two subunits form their protein-protein interactions, and these are likely to be the same regions used by the AR. Finally, the subunits cooperate to bind their element by using the inter-protein contacts so that the binding of one subunit stabilizes and enhances the effectiveness of the binding of the

second subunit. We then carried out a detailed mutagenesis study t to understand to what extent both the conserved and unique residues from a DBD confer DNA recognition (see Appendix 4, Orlowski *et al*, 2004 *in press*).

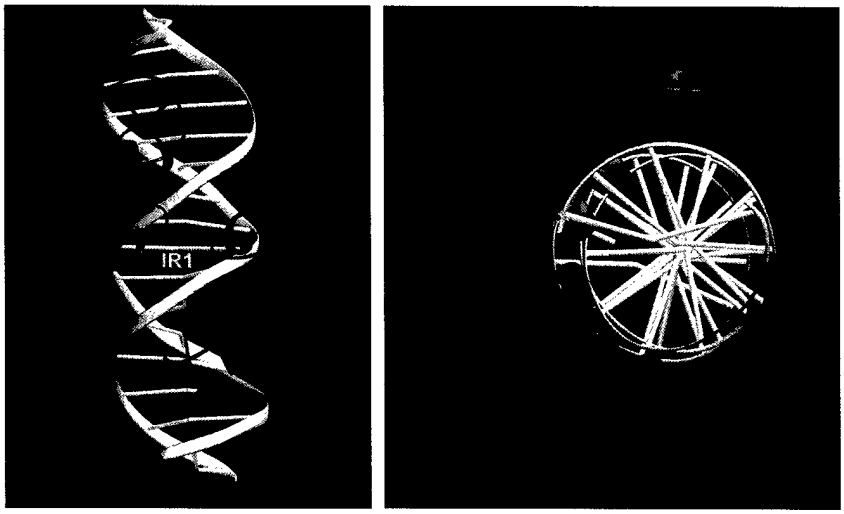



Figure 6: Two views of the EcR-Usp DNA-binding complex on their symmetric repeat DNA site. The major determinants for DNA recognition and protein dimerization are likely to be closely shared with the AR. The green spheres show the positions of the two Zn atoms, which nucleate a Zn-I and a Zn-II loop.

## Understanding additional Structural and Functional Attributes of AR

While we were carrying out our work on the DBD portion of the AR, we noticed that the some related structural work on the distinct LBD portion was reported by another group (ref 1). That study showed how androgens bind to the pocket of the AR-LBD. However, there were many unanswered questions. Among these was how the ligand alters the conformation so that the LBD can recruit p160 coactivators to bind to its surface. The binding of p160 coactivators to AR is a direct functional consequence of steroid activation and is required for AR function. While it had been known that the p160 proteins rely on discrete LXXLL motifs (where L is leucine and X is any residue) within their sequence to bind to steroid receptors, there was little structural insight on how this was actually done.

Therefore, we studied the LBD of the nuclear receptor FXR, which also binds to steroids like the AR. We recently reported (see Appendix 5, Mi *et al.*, 2003) the structure of this LBD bound to both its steroidal ligand and to p160 protein derived LXXLL motifs. In the paper, we describe the similarities with the AR as well as the mechanism of how helix-12 in these receptors can be forced in an agonist conformation that supports LXXLL sequence binding on the surface of the LBD. New and unexpected were the findings that two LXXLL motifs could bind simultaneously to the surface of the FXR by forming cooperative interactions with each other and the LBD. These results are now prompting us to examine, using a mutagenesis study, whether the AR LBD can also accommodate the binding of two p160 motifs in the presence of ligands.

Reportable Outcomes

A) Five E. coli over-expression clones for producing AR-DBD.

B) Publications and Manuscripts Pending Publication:

i) Black, B.E., Holaska, J.M, Rastinejad, F., Paschal, B.M. DNA binding domains in diverse nuclear receptors function as nuclear export signals. Current Biology, 11: 1749-1758 (2001).

ii) Holaska, J.M., Black, B.E., Rastinejad, F., Paschal, B.M. Ca2+ regulates the nuclear export activity of calreticulin. *Mol. Cell Biol.* 22:6286-97 (2002)

iii) Devarakonda, S., Harp, J.M., Kim, Y., Ozyhar, O., and Rastinejad, F. Crystal structure of the EcR-Usp DNA binding complex and its relationship to the other steroid receptors. EMBO J. 22: 5827-5840.

iv) Orlowski, M., Szyska, M., Kowalska, A., Grad, I., Zoglowek, A., Rymarczyk, G., Dobryszycki, P, Drowarsch, D., Kochman M., Rastinejad, F., Ozyhar, A. Mol. Endo. *In press* (2004)

v) Mi, L, Devarakonda, S, Harp, J.M., Han, Q., Pellicciari, R. Willson, T., Khorasanizadeh, S., Rastinejad F.

Structural basis for bile acid binding and activation of the nuclear receptor FXR.
*Molecular Cell*, 11:1093-11:00. (2003).

Conclusions:

The three years of funding plus the one additional year of no-cost extension (addendum) have seen exciting results from our laboratory in understanding the structure and function of AR domains. We have characterized an important and previously unknown function associated with the DNA-binding domain. While our structural work on the AR/DNA complex has been difficult, we were able to solve the corresponding structure of another, highly related DBD/DNA complex that shows the essential hallmarks associated with the AR protein. Finally, we have also studied the LBD regions of steroid receptors and have uncovered new mechanisms by which the AR and related members of the superfamily are activated by their steroidal ligands to bind to p160 coactivator proteins.

Appendix 1 and 2 and 3:

Publications:

Black, B.E., Holaska, J.M, Rastinejad, F., Paschal, B.M.
DNA binding domains in diverse nuclear receptors function as nuclear export signals. *Current Biology*, 11: 1749-1758 (2001).

Holaska, J.M., Black, B.E., Rastinejad, F., Paschal, B.M.
Ca2+ regulates the nuclear export activity of calreticulin.
*Mol. Cell Biol.* 22:6286-97 (2002).

Mi, L, Devarakonda, S, Harp, J.M., Han, Q., Pellicciari, R.
Willson, T., Khorasanizadeh, S., Rastinejad F.
Structural basis for bile acid binding and activation of the nuclear receptor FXR.
*Molecular Cell*, 11:1093-11:00. (2003).

Individuals Supported by this grant

Fraydoon Rastinejad, PI
Qiang Zhao (Research Associate) involved with AR-DBD/DNA crystallizaion
Christine Wright (Research Professor), working with Qiang Zhao on AR protein crystalliztation
Srikripa Devarkonda (Biophysics Graduate Student) responsible for determining the EcR-Usp Complex
LiZhi Mi (Biophysics Graduate Student) Responsible for understanding how p160 proteins bind to steroid receptor LBDs

References:

1) Matias PM, Donner P, Coelho R, Thomaz M, Peixoto C, Macedo S, Otto N, Joschko S, Scholz P, Wegg A, Basler S, Schafer M, Egner U, Carrondo MA.
Structural evidence for ligand specificity in the binding domain of the human androgen receptor. Implications for pathogenic gene mutations.
J Biol Chem. 2000: 275(34):26164-71.

# DNA binding domains in diverse nuclear receptors function as nuclear export signals

Ben E. Black*†‡, James M. Holaska*, Fraydoon Rastinejad†‡§ and Bryce M. Paschal*†‡

**Background:** The nuclear receptor superfamily of transcription factors directs gene expression through DNA sequence-specific interactions with target genes. Nuclear import of these receptors involves recognition of a nuclear localization signal (NLS) by importins, which mediate translocation into the nucleus. Nuclear receptors lack a leucine-rich nuclear export signal (NES), and export is insensitive to leptomycin B, indicating that nuclear export is not mediated by Crm1.

**Results:** We set out to define the NES in the glucocorticoid receptor (GR) and to characterize the export pathway. We found that the 69 amino acid DNA binding domain (DBD) of GR, which is unrelated to any known NES, is necessary and sufficient for export. Mutational analysis revealed that a 15 amino acid sequence between the two zinc binding loops in the GR-DBD confers nuclear export to a GFP reporter protein, and alanine-scanning mutagenesis was used to identify the residues within this sequence that are critical for export. The DBD is highly related (41%–88% identity) in steroid, nonsteroid, and orphan nuclear receptors, and we found that the DBDs from ten different nuclear receptors all function as export signals. DBD-dependent nuclear export is saturable, and prolonged nuclear localization of the GR increases its transcriptional activity.

**Conclusions:** Multiple members of the nuclear receptor superfamily use a common pathway to exit the nucleus. We propose that NLS-mediated import and DBD-mediated export define a shuttling cycle that integrates the compartmentalization and activity of nuclear receptors.

Addresses: *Center for Cell Signaling, †Department of Biochemistry and Molecular Genetics, ‡Cell and Molecular Biology Program, and §Department of Pharmacology, University of Virginia, Charlottesville, Virginia 22908, USA.

Correspondence: Bryce M. Paschal
E-mail: paschal@virginia.edu

Received: **13 August 2001**
Revised: **11 September 2001**
Accepted: **27 September 2001**

Published: **13 November 2001**

**Current Biology** 2001, 11:1749–1758

0960-9822/01/$ – see front matter


## Background

Nuclear receptors contain a structurally conserved DBD that contacts the respective DNA response element within enhancers of target genes [1–3]. Ligand-occupied nuclear receptors interact with transcriptional coactivators, whereas ligand-free nuclear receptors interact with transcriptional repressors. Ligand binding has, in addition, a profound influence on the subcellular distribution of certain nuclear receptors [4]. Glucocorticoid binding to GR induces receptor translocation from the cytoplasm to the nucleus [5], and androgen induces translocation of the androgen receptor (AR) to the nucleus as well [6]. Nuclear import occurs because ligand binding to nuclear receptors releases chaperones and exposes one or more NLSs. GR contains two ligand-regulated NLSs, one of which is located in the hinge region adjacent to the DBD [5]. The NLS is, in turn, recognized by nuclear import factors that facilitate translocation of the nuclear receptor through the nuclear pore complex and delivery to the nucleoplasm [7, 8].

Certain nuclear receptors have also been shown to undergo export from the nucleus. Hormone withdrawal from cells induces the cytoplasmic accumulation of GR and AR [6, 9]. Moreover, in the presence of ligand where nuclear receptors appear constitutively nuclear, GR, AR, thyroid hormone receptor (TR), and the progesterone receptor (PR) are rapidly shuttling between the nucleus and cytoplasm [6, 10–13]. Because the sizes of these and other nuclear receptors (43–90 kDa) exceeds the molecular weight cutoff for simple diffusion through the nuclear pore, it has been inferred that nuclear export of nuclear receptors is a facilitated process. The balance of nuclear import and export could be an important mechanism for regulating the transactivation potential of nuclear receptors. Indeed, selective nuclear localization of factors involved in signal transduction and transcription has emerged as an important regulatory step in several biological pathways [14].

The best-characterized nuclear export pathway uses Crm1 as a receptor for proteins that contain a leucine-rich NES. Nuclear receptors such as GR, AR, TR, and PR do not, however, contain a leucine-rich NES, suggesting that these proteins are not transport substrates for Crm1. This view is corroborated by the observation that the Crm1-

specific inhibitor leptomycin B does not inhibit nuclear export of GR, PR, AR, or TR in mammalian cells [9, 13, 15–17]. Recent insight into how nuclear receptors may exit the nucleus came from a biochemical screen for nuclear export factors present in extracts prepared from mammalian cells [17]. The calcium binding protein calreticulin (CRT) was identified as a factor that stimulates nuclear export of proteins that contain a leucine-rich NES [17]. The previous link between CRT and downregulation of GR, AR, all-trans retinoic acid receptor (RAR), and vitamin D receptor (VDR) transcriptional activity [18–20] suggested a potential role in the export of nuclear receptors as well, despite the absence of a leucine-rich NES in these proteins. Recombinant CRT stimulates nuclear export of GR in permeabilized cell assays, and GR export, which is deficient in *crt*$^{-/-}$ cells, can be complemented by back-transfecting CRT [17].

Here we characterize the signal for nuclear export of GR, which is contained within its DBD. Moreover, the DBD functions as the export signal in multiple steroid, nonsteroid, and orphan nuclear receptor family members. The DBD export pathway is saturable, indicating that nuclear receptors compete for a limited number of soluble transport factors, NPC binding sites, or both. Overexpression of the VDR DBD inhibits GR export and results in an elevated transcriptional response of GR to dexamethasone (Dex). Thus, the DBD export pathway contributes to the regulation of GR and probably other nuclear receptors through a nuclear transport-based mechanism.

## Results

### The DBD of GR contains the NES

The DBD of GR can be incorporated into a green fluorescent protein (GFP) reporter and can function as an NES when expressed in cultured cells [17]. The conserved sequence and structure of the DBD in different nuclear receptors (Figure 1a,b) led us to hypothesize that the DBD is widely used as an export signal. To address this hypothesis, we first mapped the molecular determinants of the 69 amino acid GR DBD that are necessary for nuclear export by using a GFP reporter that undergoes ligand-dependent nuclear import (Figure 1c). Following ligand removal from the cell culture media, the distribution of the GFP reporter was recorded at 0, 2, 4, and 6 hr by fluorescence microscopy. We found that cysteine to alanine mutations that disrupt the zinc binding loops and abolish DNA binding only slightly inhibited nuclear export of the GFP reporter (C424A and C463A; Figure 2). This suggested that nuclear export might instead rely on the DNA recognition helix that is situated between the two zinc binding loops. Indeed, a GFP reporter that contained only 15 amino acids of the GR DBD (442–456) underwent translocation from the nucleus to the cytoplasm. Although the zinc binding loops of nuclear receptors do not contain information that is critical for export specified by the DBD, there may be a contribution to the DBD structure that augments recognition of the signal by the export machinery.

Alanine-scanning mutagenesis of the DNA recognition helix between the loops revealed that a pair of phenylalanines is critical for the export function of the DBD (FF→AA; Figure 2). Mutating other pairs of amino acids between the zinc fingers (KR→AA, VE→AA, and YL→AA) impaired the nuclear export function of the GR DBD to a lesser degree (2 and 4 hr time points). In contrast, other mutations (KV→AA, GQ→AA, and HN→AA) had no effect on the export activity of the GR DBD. Our data show that amino acids between the two zinc binding loops, which include part of the DNA recognition helix, contain the NES of GR.

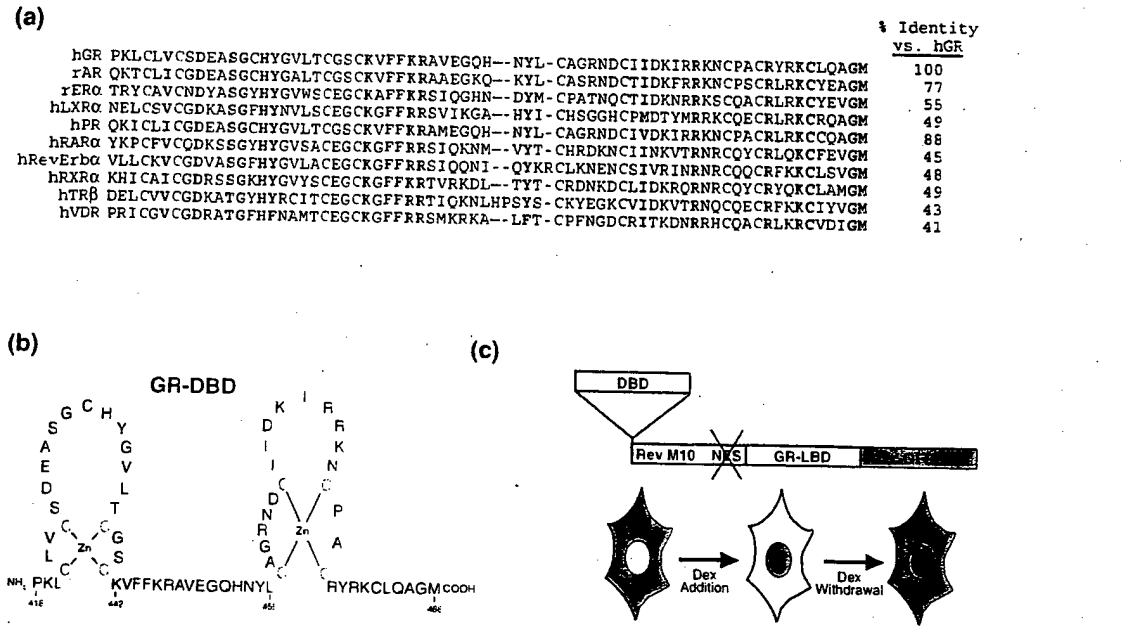
### Steroid, nonsteroid, and orphan receptors use a common export pathway

Our finding that the GR DBD contains the NES appeared significant in light of the structural conservation of the DBD among nuclear receptors. The GR DBD is at least 40% identical to the sequences of DBDs from nine different nuclear receptors included in our analysis (Figure 1a). Moreover, every member of the nuclear receptor superfamily that has been identified to date contains a pair of phenylalanines that are located at the same position relative to phenylalanines within the GR DBD. This led us to hypothesize that the DBDs from diverse nuclear receptors can function as an NES. We tested this hypothesis by analyzing the export activity of DBDs from steroid, nonsteroid, and orphan nuclear receptors. We found that nuclear export of the GFP reporter was conferred by the DBD derived from AR, estrogen receptor (ER), liver X receptor (LXR), PR, RAR, RevErb, retinoid X receptor (RXR), TR, and VDR (Figure 3). Nuclear export was not conferred by the DBD derived from GATA1, a transcription factor that, like the nuclear receptors, contains two zinc binding loops in its DBD [21]. GATA1 is not, however, a member of the nuclear receptor superfamily based on sequence or structure. Thus, our data identify the DBD of nuclear receptors as a new type of NES. The DBD does not contain a leucine-rich NES that is recognized by the export receptor Crm1, which may explain why nuclear export of GR, PR, TR, and AR are insensitive to leptomycin B [9, 13, 15–17].

### Mutations in the DBD block nucleocytoplasmic shuttling of nuclear receptors

Our experiments demonstrated that the DBD is sufficient to mediate nuclear export in the context of the GFP reporter. To determine whether the DBD is necessary for export in the context of the receptor, we examined nucleocytoplasmic shuttling of full-length, wild-type (wt) GR, AR, and RAR, and of mutant forms of these receptors that were predicted to be deficient for export. We designed a modified-shuttling assay that, in a heterokaryon

**Figure 1**



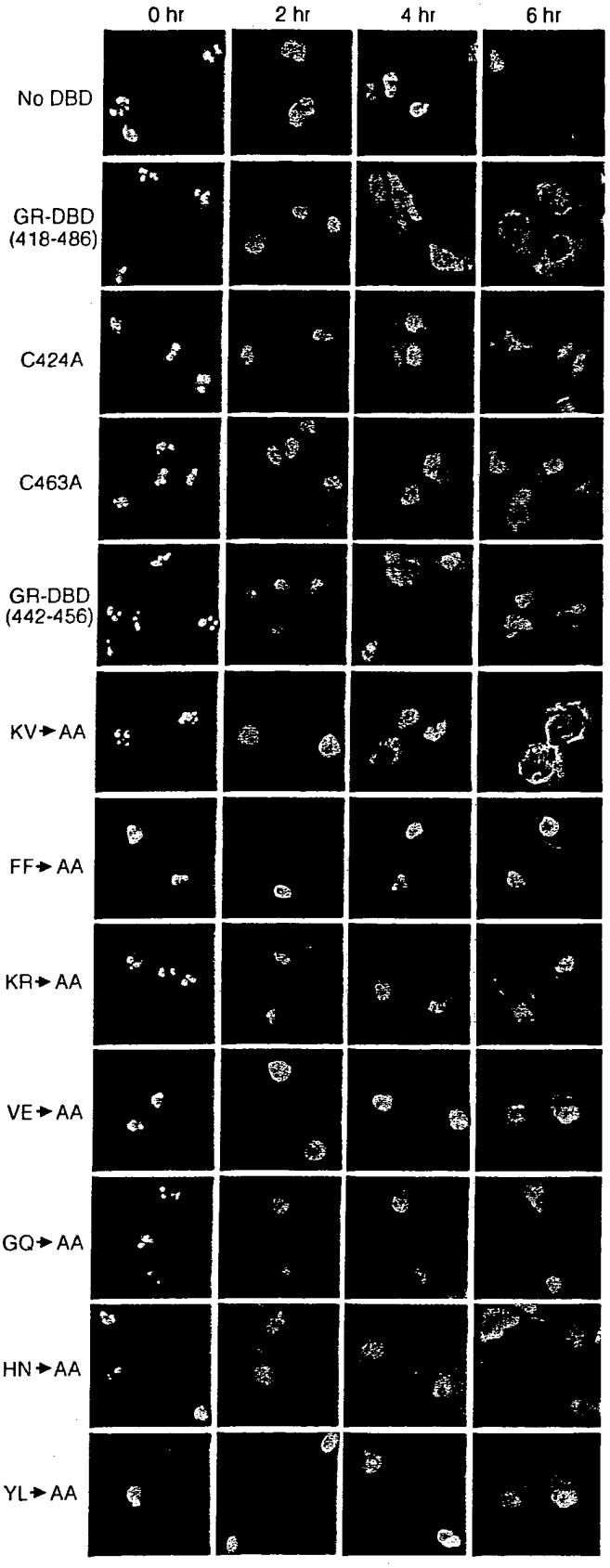

Conserved structure of the DBD and assay for nuclear export activity. **(a)** Alignment of DBDs from nuclear receptors used in this study and the percent identity to the DBD of human GR. Highly conserved residues (bold) including the cysteines that coordinate zinc binding (green) and the pair of phenylalanines that are present in the DNA recognition helix of all nuclear receptors (red) are indicated. **(b)** Diagram of the GR DBD showing the position of the pair of phenylalanines relative to the zinc binding loops. **(c)** Features of the GFP reporter and two-step assay for nuclear export in vivo. The vector encodes a mutant form of Rev (M10) that contains a nonfunctional NES fused to the ligand binding domain (LBD) of GR and GFP. Dex induces nuclear import of the GFP reporter, which remains nuclear after Dex withdrawal because the reporter lacks a functional NES. In-frame fusion of a DBD to the N terminus of the GFP reporter restores nuclear export activity.

fusion, scores nuclear export from a donor nucleus and nuclear import into an acceptor nucleus. Prior to fusion (Figure 4a, upper panels), the donor cells (human) were transfected with GFP fusions of the receptors and treated with the appropriate ligand to induce nuclear import. The acceptor cells (mouse) were labeled with a CellTracker dye that, following its uptake, is converted to a membrane impermeant form in the cytoplasm. After fusion (Figure 4a, lower panels), the cytoplasm of the heterokaryon fluoresces red, and the GFP fusion (green) containing full-length GR equilibrates between the human and mouse nuclei because it undergoes shuttling, as previously shown [11]. The GFP reporter used in prior experiments to characterize the DBD (Figure 1c) undergoes nuclear import, but fails to undergo export and equilibrate between nuclei in the heterokaryon because it lacks a functional NES. Nuclear export of the GFP reporter was restored by the GR DBD, resulting in equilibration of the reporter between nuclei in the heterokaryon fusion (Figure 4b).

The GFP reporter used in the experiments described above contains the ligand binding domain of GR to direct Dex-dependent import. We used another GFP reporter, which lacked any GR sequence, to formally test for the sufficiency of the DBD export signal. The GFP reporter used contains the SV40 large T antigen NLS to direct constitutive import, but it remains nuclear in the heterokaryon because it does not contain a nuclear export signal (Figure 4c, upper panels). Fusion of the GR DBD to the reporter resulted in nucleocytoplasmic shuttling (Figure 4c, lower panels). Thus, the GR DBD is sufficient to direct nuclear export of the GFP reporter.

We next transfected full-length nuclear receptors (GR, AR, and RAR) containing either a wt DBD or a mutant DBD (FF→AA) into donor cells, treated the cells with ligand, and fused them with acceptor cells labeled with the CellTracker dye. The wt forms of GR, AR, and RAR equilibrated between nuclei in the heterokaryon fusion (Figure 4d), indicating that export of these nuclear receptors occurs even in the presence of ligand. The DBD mutant forms (FF→AA) of GR and AR, however, failed to equilibrate between nuclei in the heterokaryon fusion. Thus, export of GR and AR from the donor nuclei is inhibited by changing only two amino acids within the DBD of each nuclear receptor. Surprisingly, the DBD mutant form of RAR still equilibrated between nuclei in the heterokaryon fusion (Figure 4d). Because the DBD of RAR is sufficient to specify nuclear export of the GFP reporter used in our analysis (Figure 2), we interpret this

**Figure 2**




as evidence that RAR export can occur by an alternative, DBD-independent transport pathway. Nuclear export of RXR in nerve cells can be facilitated by its binding partner NGFI-B, an orphan nuclear receptor with three leucine-rich NESs [22]. Nuclear export of the wild-type and mutant forms of RAR, however, still occurs in the presence of leptomycin B (data not shown). Certain nuclear receptors may, therefore, be exported from the nucleus by more than one transport pathway.

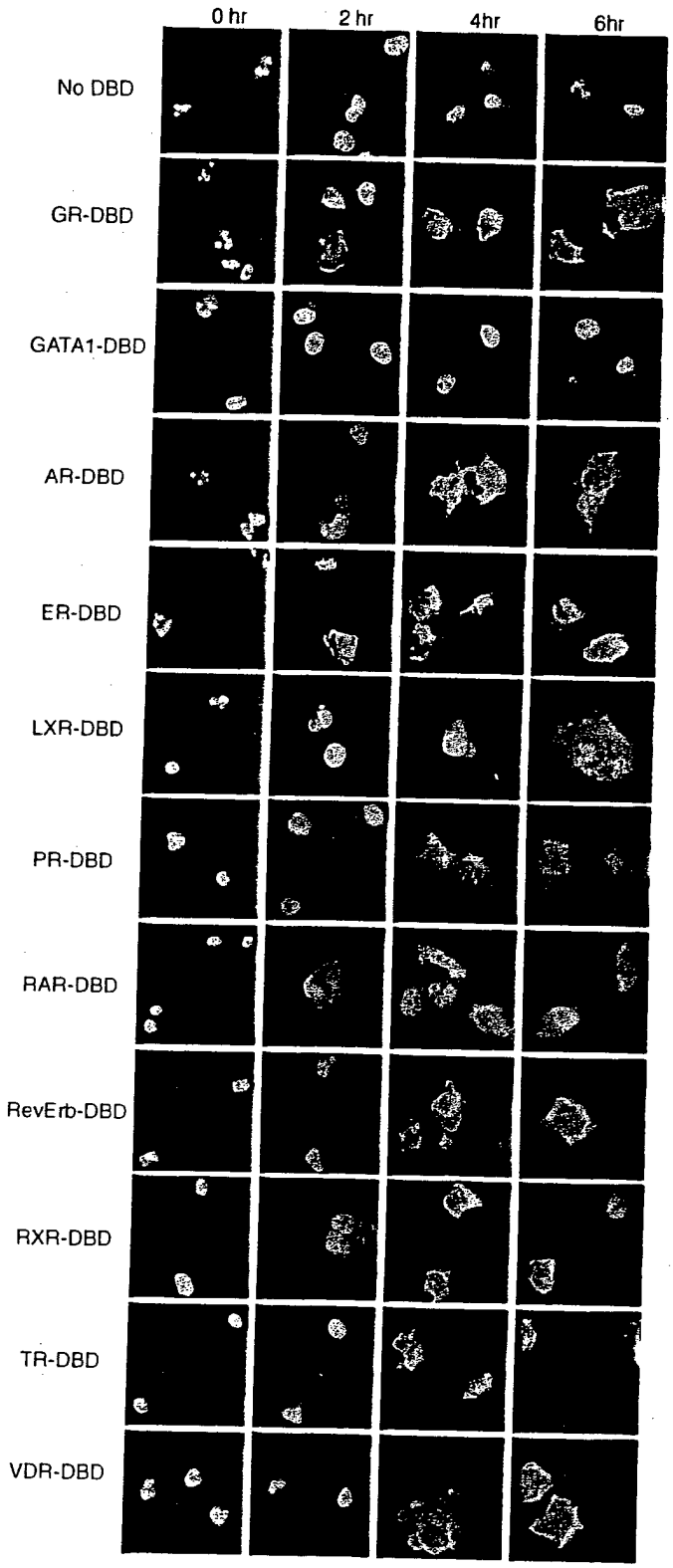
**CRT binding to the GR DBD is required for nuclear export**

Our results showing that the GR DBD functions as an export signal together with our previous finding that GR export is defective in *crt*$^{-/-}$ cells suggested a molecular mechanism for the transport pathway [17]. This could involve CRT-dependent recognition of the GR DBD in the nucleus and translocation of the complex to the cytoplasm. We reasoned that if CRT binding to the DBD is an obligate step in GR export, then mutations within the DBD that inhibit export might also inhibit the interaction of GR with CRT. We tested this by using binding assays with recombinant CRT and glutathione-*S*-transferase (GST) fusions of the wt and mutant (FF→AA) forms of the GR DBD. CRT bound to the wt DBD but displayed only background binding to the GR DBD mutant protein (Figure 5a). Crm1 did not bind to the GR DBD, consistent with other data indicating that Crm1 is not the receptor for nuclear export of GR [16, 17].

The net redistribution of GR from the nucleus to the cytoplasm generally requires several hours in ligand-free media; however, the kinetics of export can enhanced by microinjecting recombinant CRT into cells [17]. We used this approach to examine whether GR containing a wt or mutant DBD would undergo CRT-dependent nuclear export. Cos cells were transfected with GFP fused to full-length GR (wt or FF→AA mutant) and treated with ligand to induce nuclear import of GR. The cells were then injected with CRT, and the distribution of GR was recorded by fluorescence microscopy 30 min postinjection.

The DBD of GR functions as an NES. The GFP reporter engineered with the DBD of GR was transfected into Cos cells and assayed for nuclear export. Images of living cells were recorded by fluorescence microscopy at 0, 2, 4, and 6 hr following Dex withdrawal. The GFP reporter alone (no DBD) remains nuclear during the course of the experiment, while including the GR DBD (residues 418–486) in the GFP reporter confers export by 4 hr. Point mutations in the GR DBD (C424A or C463A) slightly reduced the level of export, evident at 4 hr. The 15 residue sequence (KVFFKRAVEGQHNYL; residues 442–456) that resides between the two zinc binding loops was sufficient for export of the GFP reporter. Alanine-scanning mutagenesis of residue pairs in the DBD revealed that the FF→AA mutation in the DNA recognition helix causes a major defect in nuclear export. The KR→AA, VE→AA, and YL→AA mutations have an intermediate effect on nuclear export, and the KV→AA, GQ→AA, and HN→AA mutations have no effect on nuclear export.

**Figure 3**




The DBDs of diverse nuclear receptors can all function as an NES. The DBDs from hormone, nonhormone, and orphan nuclear receptors were analyzed for nuclear export activity in vivo by using the GFP reporter. The GFP reporter alone (no DBD) remains nuclear during the course of the experiment. The DBD from the transcription factor GATA1, which is similar in size to the nuclear receptors and contains two zinc binding loops, does not confer nuclear export to the GFP reporter. The DBDs from nine different nuclear receptors (AR, ER, LXR, PR, RAR, RevErb, RXR, TR, and VDR) confer nuclear export activity to the GFP reporter.

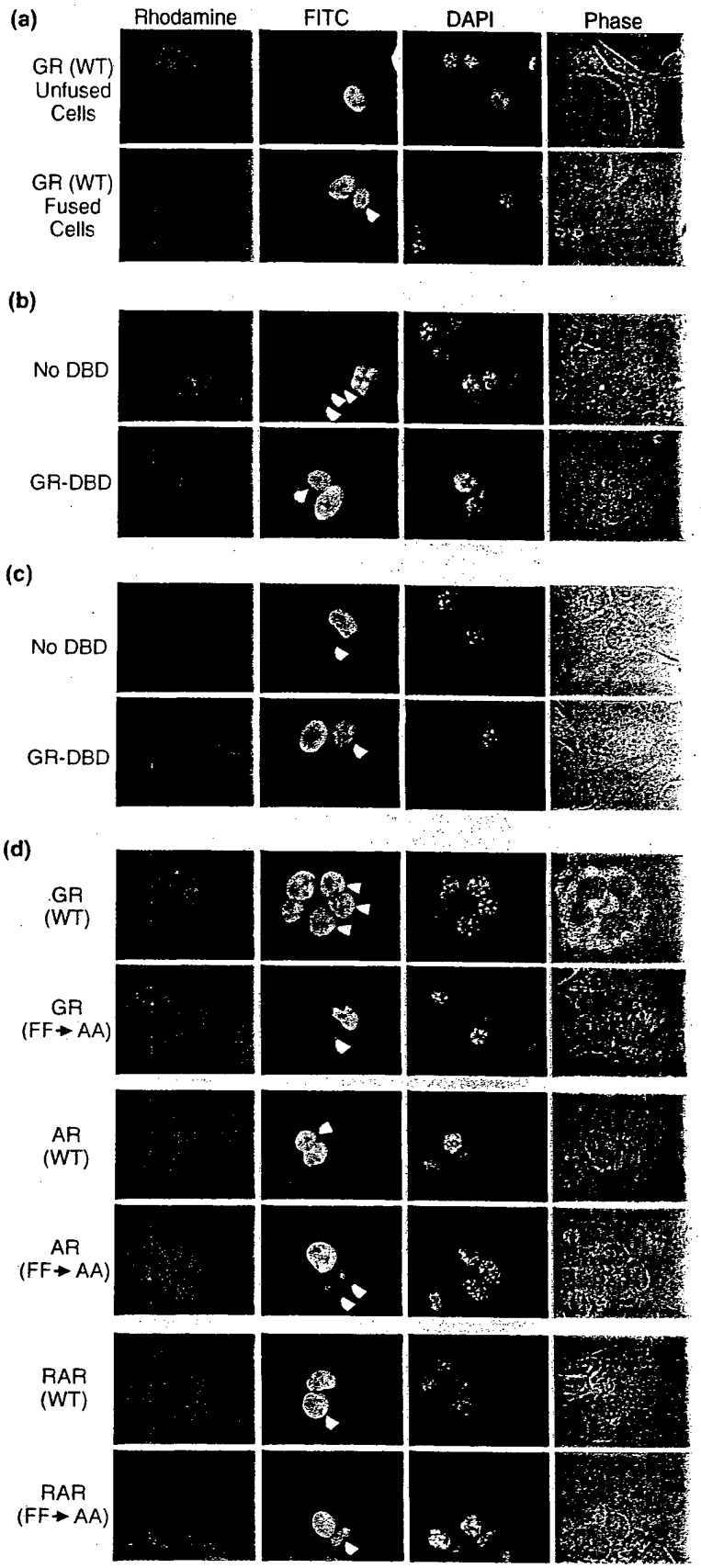
We observed that injection of cells with CRT caused the relocalization of wt GR from the nucleus to the cytoplasm (Figure 5b, upper panels). In contrast, CRT injection into cells expressing the GR DBD mutant (FF→AA) did not affect GR distribution (Figure 5b, middle panels). Crm1 injection did not cause wt GR redistribution to the cytoplasm. Thus, a functional export signal within the DBD is necessary for CRT-dependent translocation of GR from the nucleus to the cytoplasm. Recombinant CRT can also mediate nuclear export of GR in a permeabilized cell assay [17], and this can be inhibited by the addition of the DBD from VDR (Figure 5c). The ability of the DBD from VDR to act as a competitive inhibitor of GR export indicates that CRT recognizes similar molecular determinants in the DBDs of different nuclear receptors.

### Excess DBD blocks GR export and increases transcription

Our finding that the VDR DBD can act as a competitive inhibitor of GR export in permeabilized cells led us to examine whether overexpression of the VDR DBD fused to BFP would compete with the GR export pathway in living cells. Because the DBD of VDR also contains an NLS [23], it undergoes export and import but is concentrated in the nucleus at steady state. Expression of BFP alone did not affect nuclear export mediated by the GR-DBD examined at 0 and 5 hr (Figure 6, upper panels). In contrast, expression of BFP-VDR DBD inhibited GR-DBD export and resulted in nuclear accumulation of GR-DBD (Figure 6, middle panels). This effect required a functional DBD in VDR since mutating the conserved pair of phenylalanines in the DBD reversed the inhibition (Figure 6, lower panels). Our data support the hypothesis that a common pathway is used for nuclear export of both steroid (GR) and nonsteroid (VDR) nuclear receptors.

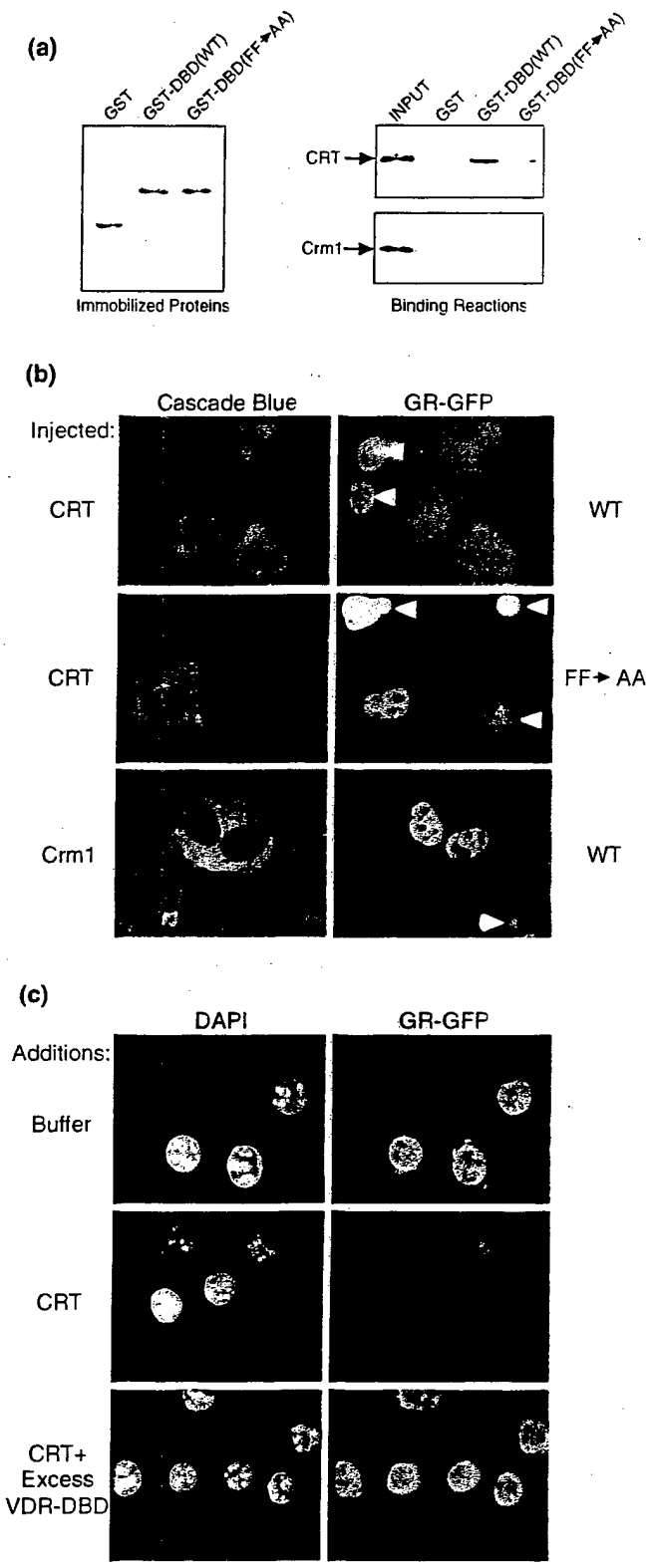
Inhibiting DBD-dependent export in vivo restricts GR to the nucleus, thereby blocking the nucleocytoplasmic shuttling cycle. This allowed us to examine whether blocking nuclear export could augment the GR-dependent transcription by increasing its dwell time in the nucleus. We also considered the possibility that nucleocytoplasmic shuttling could be necessary for the transcriptional activity of GR because, for example, cytoplasmic chaperones are required for ligand binding. GR-dependent transcription can be measured in the presence of excess VDR DBD because VDR does not bind to a GRE [24]. The response elements for GR and VDR differ in sequence, direction, and spacing of the half-sites.

**Figure 4**




The DBD export signal is necessary for hormone receptor shuttling in vivo. **(a)** Heterokaryon shuttling assays were performed with Cos cells transfected with full-length GR fused to GFP (FITC) and NIH 3T3 cells labeled with the dye CellTracker CMTMR (Rhodamine). When coseeded on coverslips and fused by brief (30 s) incubation in polyethylene glycol (Roche; 50% vol:vol), the Cos and 3T3 cells fuse and fluoresce red. Nucleocytoplasmic shuttling (export and import) of the GFP reporter results in equilibration of green fluorescence between the donor Cos cell nuclei and the acceptor 3T3 cell nuclei within the heterokaryon. Acceptor cell nuclei (white arrowheads) are also distinguished by centromeric foci that stained brightly with DAPI. **(b)** The GR DBD restores shuttling behavior to the GFP reporter that contains a nonfunctional NES. **(c)** The GR DBD confers shuttling activity to a GFP reporter that contains the SV40 large T antigen NLS [38]. The GFP reporter contains the streptavidin gene and assembles into a ~160 kDa tetramer that is too large to escape the nucleus by diffusion. **(d)** The DBD export signal is necessary for nucleocytoplasmic shuttling in the context of full-length hormone receptors. Full-length wt or mutant (FF→AA) GR, AR, and RAR were tested for nucleocytoplasmic shuttling in the presence of 1 μM Dex, 10 nM R1881 (synthetic androgen), or 1 μM all-trans retinoic acid, respectively. Each wt receptor equilibrates between the nuclei of a heterokaryon. The FF→AA mutation inhibits GR and AR shuttling, whereas the FF→AA mutation has only a slight effect on RAR shuttling. The FF→AA mutation in the GR DBD abolishes DNA binding in vitro, suggesting the mutation does not generate a nuclear retention signal (data not shown).

**Figure 5**



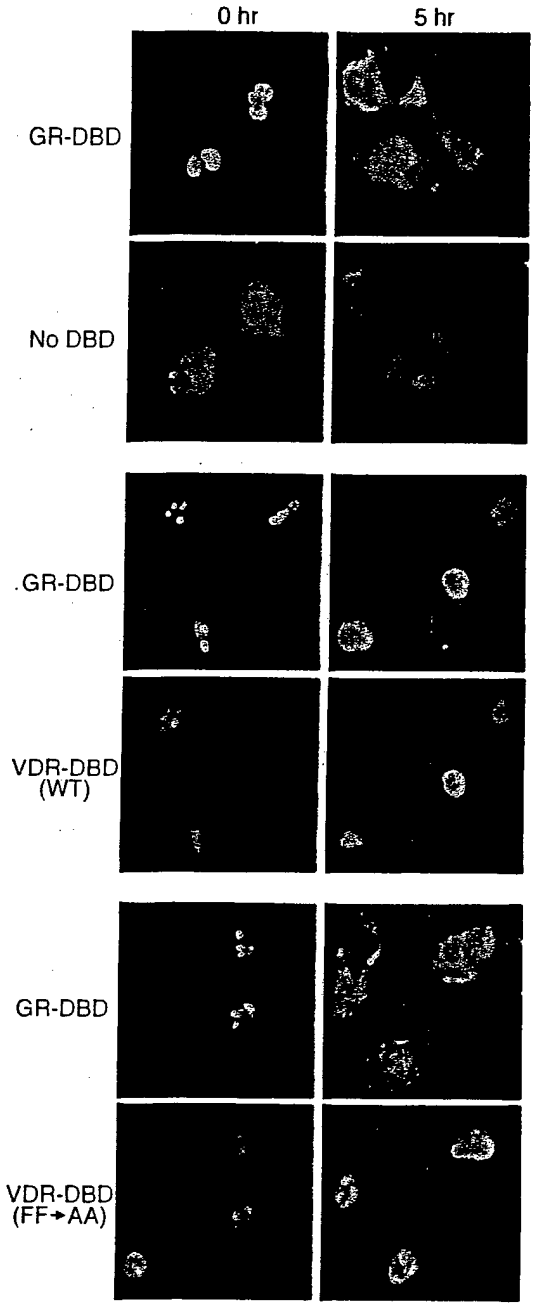

The DBD export signal is necessary for export mediated by CRT. **(a)** Direct binding of CRT to the GR DBD. GST fusion proteins containing the GR DBD (wt or FF→AA) were immobilized on glutathione-Sepharose beads and used for binding assays with recombinant CRT and Crm1. CRT binds to the wt DBD from GR but not to the export-defective mutant DBD (FF→AA) from GR. A preparation of recombinant Crm1 that binds the leucine-rich NES [39] does not bind to the DBD from GR. Shown are 5% of the total input proteins (CRT or Crm1). **(b)** CRT stimulates nuclear export of wt but not the export-defective mutant of GR in living cells. Cos cells transfected with GR (wt or FF→AA mutant) fused to GFP were treated with Dex to induce nuclear import. Following Dex withdrawal, the cells were microinjected with either His-tagged CRT or His-tagged Crm1 (1 mg/ml each), and the distribution of GR-GFP was examined after 30 min. A 70 kDa fluorescent dextran (Cascade Blue) was included to mark the injection site and to verify that the nuclear envelope remained intact during the experiment. CRT-dependent export of wt GR was observed in 30/30 cells injected, whereas CRT-dependent export of the mutant GR (FF→AA) was observed in only 2/21 cells injected (three experiments). GR did not undergo nuclear export in uninjected cells (white arrowheads) or Crm1-injected cells in this assay. The latter control is consistent with Crm1-independent export of GR [16, 17]. **(c)** Shared mechanism for nuclear export of GR and VDR. A cell line expressing GR-GFP [35] was permeabilized with digitonin and incubated with recombinant CRT (100 μg/ml) in the absence and presence of the DBD from VDR (120 μg/ml). CRT-dependent export of GR-GFP is blocked by the VDR DBD.

GR-dependent transactivation of a GRE-luciferase reporter was measured in the presence of cotransfected BFP or BFP-VDR DBD as a function of Dex concentration (Figure 7a). Expression of the BFP-VDR DBD increased GRE luciferase activity 2-fold relative to BFP. Thus, inhibiting GR export results in an increase in GR-dependent transcription. The effect was correlated with the amount of BFP-VDR DBD plasmid used for cotransfection (Figure 7b). We also found that a mutant form of the VDR DBD (FF→AA) that has reduced activity as a competitive inhibitor of GR export (Figure 6) increases the transcriptional activity of GR only slightly (Figure 7c). Our data indicate that nuclear export of GR is critical for proper regulation of its transcriptional activity, since blocking GR export results in an elevated response to ligand. Moreover, nuclear export is predicted to be important for regulating the cytoplasmic functions that are emerging for certain nuclear receptors.

## Discussion

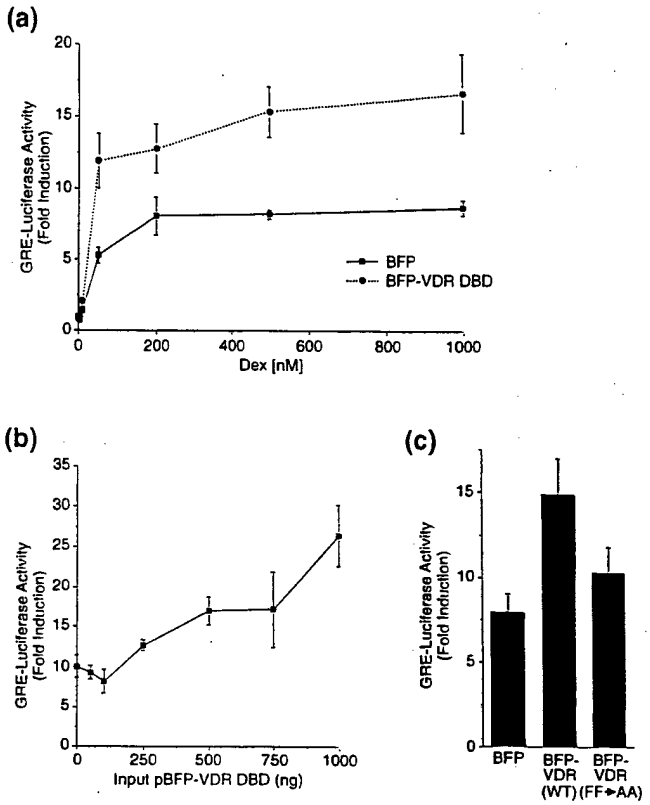
In the context of a living cell, GR undergoes rapid exchange between chromatin binding sites and the nucleoplasm [25]. Upon dissociating from chromatin, GR may receive input from signaling pathways that regulate its transcriptional activity, either in the form of posttranslational modifications or cofactor interactions [26]. This could occur in the nucleoplasm, or it could occur during the movement of GR through the cytoplasm during its shuttling cycle. The latter scenario would obviate the requirement for nuclear import of cytoplasmic enzymes or cofactors, possibly to ensure spatial separation from the nuclear receptors. Because GR, AR, and RAR all shuttle in the presence of their respective ligand, nucleocytoplasmic shuttling is predicted to be a general property of nuclear

**Figure 6**




Competition for DBD-mediated export in vivo. The GFP reporter engineered with the DBD of GR was cotransfected with a plasmid encoding BFP alone (no DBD) or BFP fused to the DBD of VDR in Cos cells. Images were captured at 0 and 5 hr following Dex withdrawal. BFP alone did not alter import of the reporter alone (data not shown) or the import and export of the reporter containing the DBD of GR. The DBD of VDR (VDR-DBD [wt]) blocks GR DBD-mediated nuclear export. However, a mutant VDR DBD that has two critical phenylalanines replaced with alanines (VDR-DBD [FF→AA]) does not compete for GR DBD-mediated nuclear export.

**Figure 7**




Blocking GR export increases ligand-induced gene expression. **(a)** VDR-DBD expression increases GR-mediated gene expression over a wide range of Dex concentration. Endogenous GR activity was examined in NIH3T3 cells using a GRE-luciferase reporter. Cells were cotransfected in 6 well dishes with plasmids encoding a *Renilla*-luciferase reporter (30 ng), GRE-firefly luciferase (370 ng), and either BFP alone or BFP-VDR DBD (1000 ng), and then they were incubated overnight. Cells were lysed 6 hr following the addition of the indicated amount of Dex. **(b)** Increasing input pBFP-VDR DBD increases Dex-induced gene expression. Transfectants were incubated overnight and then treated with Dex (500 nM) for 6 hr prior to lysis. **(c)** Expression of a VDR DBD mutant that fails to compete for GR DBD-mediated nuclear export only slightly increases Dex-induced gene expression. Plasmids encoding either wt or FF→AA mutant versions of BFP-VDR DBD (500 ng each) were transfected and assayed as above. All luciferase assays were normalized for transfection efficiency using *Renilla*-luciferase expression. The values shown represent the fold induction by the addition of Dex and are the averages of triplicate wells plus or minus the standard deviation.

receptors. This would integrate cytoplasmic signaling events with regulation of nuclear transcription [27]. Alternatively, the function of DBD-dependent export may relate to cytoplasmic functions for nuclear receptors that may be transcription-independent [28–30].

Our finding that the DBD of nuclear receptors performs dual targeting functions in the cell is not without precedent. The DBD in Gal4 contains an NLS, and the DBD in Stat1 contains a leucine-rich NES that targets it for export by the Crm1 pathway [31, 32]. Thus, protein do-

mains that bind DNA provide suitable structures for presenting signals to the nuclear transport machinery. Nuclear transport receptor binding to a DBD-containing protein should, in addition, block the interaction of the protein with DNA. In the case of GR and probably other nuclear receptors, this may impart negative regulation on transcription even prior to nuclear export [18–20]. Nuclear export facilitated by DBDs and nuclear import facilitated by ligand-regulated NLSs illustrate how simple nuclear transport-based mechanisms are used to regulate the activity of a superfamily of transcriptional activators and influence gene expression in multiple biological pathways.

## Materials and methods

### *Nuclear receptor cDNAs*

The DBD from human GR was subcloned from the previously described plasmid pK7-GR-GFP [33]. The GenBank accession numbers corresponding to the other nuclear receptor superfamily members used for the DBD analysis are: rat AR, M23264; rat ERα, NM_012689; human LXRα, NM_005693; human PR, NM_000926; human RARα, NM_000964; human RevErbAα, X72631; human RXRα, NM_002957; human TRβ, X04707; and human VDR, NM_000376. Mutations were made using the QuickChange system (Stratagene).

### *Nuclear export assays*

In vivo export assays

The reporter used to test DBD-mediated nuclear export has been described [17] and is derived from the plasmid pXM10 [34]. Cos7 cells were seeded onto coverslips and grown overnight prior to transfection of the reporter vectors using the transfection reagent Fugene 6 (Roche). All BFP-DBD competition experiments used plasmids derived from pEBFP-C (Clontech). The transfectants were grown overnight prior to addition of Dex (1 μM) for 45–60 min. Cells were washed five times with serum-free and phenol red-free Dulbecco's modified eagle medium (DMEM) and incubated for the indicated times in phenol red-free DMEM containing 10 μg/ml cyclohexamide and 10% charcoal stripped newborn calf or fetal bovine serum. Cells were fixed in formaldehyde (3.7%) and processed for fluorescence microscopy. Digital images were captured by a charge-coupled device camera (Hamamatsu ORCA) mounted on a Nikon Microphot-SA microscope, with Openlab (version 2.0.6) software. Figures were assembled with Adobe Photoshop (version 5.5) and Freehand (version 9). The examples shown are representative of 20–50 cells observed in two or more independent experiments.

In vitro export assays

A cell line expressing GR-GFP [35] was treated with 1 μM Dex for 45 min. Cells were permeabilized with 0.005% digitonin for 5.5 min. Export reactions were performed at 30°C for 30 min in transport buffer (20 mM Hepes [pH 7.4], 110 mM potassium acetate, 2 mM magnesium acetate, and 1 mM EGTA) supplemented with an ATP-regeneration system, 2 mM dithiothreitol (DTT), and a protease inhibitor cocktail including aprotinin, leupeptin, and pepstatin (each at 1 μg/ml). GFP images were captured with the same exposure times.

### *Heterokaryon analysis*

The modified interspecies heterokaryon analysis was performed as follows. NIH3T3 cells were labeled in tissue culture dishes with 500 nM CellTracker dye, (5 (and 6)-(((4-chloromethyl) benzoyl) amino) tetramethylrhodamine (CMTMR; Molecular Probes), according to the manufacturers instructions. Unincorporated dye was removed and cells were trypsinized and coseeded on glass coverslips with Cos7 cells that had been transfected with the indicated plasmids using the transfection reagent Fugene 6 (Roche). Equal numbers of each cell type were seeded for a total of $3 \times 10^5$ cells per 35 mm dish and grown overnight prior to fusion. The indicated ligands were added to the cells 45 min prior to fusion, in order to induce nuclear accumulation of the reporter proteins. Cells were incubated at 37°C for 4 hr postfusion in the presence of the indicated nuclear receptor ligands and 10 μg/ml cyclohexamide to block protein synthesis. Cells were then fixed and processed for fluorescence microscopy. AR was expressed with an N-terminal FLAG epitope and detected by indirect immunofluorescence using the monoclonal antibody M2 (Sigma) at a dilution of 1:2000. Goat anti-mouse fluorescein conjugated secondary antibody (Pierce) was used at a dilution of 1:100. GR and RAR were expressed as GFP fusions and detected in the FITC channel. The examples shown are representative of 10–25 heterokaryons from two or more independent experiments.

### *Microinjection analysis*

Cos7 cells were transfected with the indicated vectors and grown overnight in 60 mM dishes; $1.2 \times 10^5$ cells per 35 mM dish were seeded onto gridded coverslips and grown overnight. Cells were treated with 1 μM Dex for 45 min to induce nuclear import of the reporter proteins. Following Dex withdrawal, cells were microinjected with the indicated transport factors and a cascade blue injection marker (1 mg/ml; MW = 70 kDa) using femptotips mounted on a Micromanipulator and Transjector (Eppendorf). Cells were incubated 30 min at 37°C and then fixed and processed for fluorescence microscopy.

### *Recombinant proteins*

Mouse CRT was subcloned in pQE-30 (Qiagen) for expression as a His-tagged protein in the TG1 *E. coli* strain. Cultures were grown in Luria Broth for 24 hr at 37°C without induction, and His-CRT was purified on TALON beads (Clontech), eluted with imidazole, and dialyzed into 50 mM Hepes (pH 7.4). The expression and purification of His-Crm1 has been described [36]. Recombinant transport factors were flash-frozen in single-use aliquots and stored at −80°C. The nuclear receptor DBDs were expressed in *E. coli* as GST fusions by using pGEX-2T and pGEX-4T vectors (Pharmacia). Cells were grown in Luria Broth, and protein expression was induced when $OD_{600}$ reached 0.5, with 1 mM IPTG. The cells were collected by centrifugation, resuspended in lysis buffer (25 mM Tris [pH 7.5] and 150 mM sodium chloride), and sonicated at 5°C until lysis was complete. The lysate was centrifuged at 15,000 × g for 60 min, and the supernatant was loaded onto a glutathione-Sepharose column. The column was equilibrated and washed extensively with the lysis buffer, and GST-DBDs were eluted with the same buffer supplemented with 10 mM fresh glutathione. The GST-DBD proteins, in some cases, were cleaved with thrombin (Sigma) to generate the free DBDs. The GST and thrombin were removed using a strong cation exchange media SP-Sepharose (Pharmacia). The column was developed with 25 mM Tris buffer and a 150–500 mM sodium chloride gradient, with the GST and thrombin eluting at the beginning of the gradient and the free DBDs eluting at 300–350 mM sodium chloride.

### *Binding assays*

His-CRT or His-Crm1 (each at 1 μg/ml) proteins were incubated for 30 min at 4°C with the indicated GST proteins immobilized on glutathione-Sepharose beads. Binding reactions were carried out in 0.5X transport buffer supplemented with 10 mg/ml bovine serum albumin, 2 mM DTT, and 0.2% Tween-20. Beads were subsequently washed three times, and bound proteins were eluted with SDS-PAGE sample buffer and analyzed by Western blotting. CRT was detected using an anti-CRT rabbit polyclonal antibody (Stressgen), and Crm1 was detected using an anti-Crm1 rabbit polyclonal antibody (kindly provided by Ralph Kehlenbach, Scripps) [37].

### *Luciferase assays*

Transcriptional reporter activity was measured by using the Dual-Luciferase Reporter Assay System (Promega) according to the manufacturer's directions. NIH3T3 cells were seeded at $2 \times 10^5$ per well of a 6 well dish and grown overnight prior to transfection. Transfections of the indicated plasmids were performed with the Cytofectene transfection reagent (Bio Rad). The Renilla-luciferase (R-Luc; Promega) and GRE-luciferase (generously provided by Gordon Hager, NIH) reporter vectors were used in all cases. The total input DNA was 1.4 μg per well in

each experiment. Transfectants were grown overnight in phenol red-free DMEM containing 10% charcoal stripped fetal bovine serum prior to the indicated Dex treatments. Cells were washed with phosphate-buffered saline and lysed with 300 µl passive lysis buffer (Promega). Luciferase activities were measured by using a Berthold LB 953 luminometer.

## Acknowledgements

We thank our University of Virginia colleagues D. Allis, D. Burke, B. Pearson, L. Pemberton, and D. Wotton for reading the manuscript. We also thank G. Hager, J. Hanover, I. Macara, M. Weber, D. Wojchowski, and D. Wotton for their kind gifts of reagents, and M. Skernolis and N. Mou for technical assistance. Financial support from the National Institutes of Health (GM58639-01 to B.M.P.) and the Department of Defense (DAMD 17-00-1-0048 to F.R.) is gratefully acknowledged.

## References

1. Evans RM: **The steroid and thyroid hormone receptor superfamily.** *Science* 1988, **240:**889-895.
2. Mangelsdorf DJ, Thummel C, Beato M, Herrlich P, Schütz G, Umesono K, *et al.*: **The nuclear receptor superfamily: the second decade.** *Cell* 1995, **83:**835-839.
3. McKenna NJ, Lanz RB, O'Malley BW: **Nuclear receptor coregulators: cellular and molecular biology.** *Endocrinol Rev* 1999, **20:**321-344.
4. Hager GL, Lim CS, Elbi C, Baumann CT: **Trafficking of nuclear receptors in living cells.** *J Steroid Biochem Mol Biol* 2000, **74:**249-254.
5. Picard D, Yamamoto KR: **Two signals mediate hormone-dependent nuclear localization of the glucocorticoid receptor.** *EMBO J* 1987, **6:**3333-3340.
6. Georget V, Lobaccaro JM, Terouanne B, Mangeat P, Nicolas JC, Sultan C: **Trafficking of the androgen receptor in living cells with fused green fluorescent protein-andrógen receptor.** *Mol Cell Endocrinol* 1997, **129:**17-26.
7. Nakielny S, Dreyfuss G: **Transport of proteins and RNAs in and out of the nucleus.** *Cell* 1999, **99:**677-690.
8. Wente SR: **Gatekeepers of the nucleus.** *Science* 2000, **288:**1374-1377.
9. Tyagi RK, Lavrovsky Y, Ahn SC, Song CS, Chatterjee B, Roy AK: **Dynamics of intracellular movement and nucleocytoplasmic recycling of the ligand-activated androgen receptor in living cells.** *Mol Endocrinol* 2000, **14:**1162-1174.
10. Guiochon-Mantel A, Lescop P, Christin-Maitre S, Loosfelt H, Perrot-Applanat M, Milgrom E: **Nucleocytoplasmic shuttling of the progesterone receptor.** *EMBO J* 1991, **10:**3851-3859.
11. Madan AP, DeFranco DB: **Bidirectional transport of glucocorticoid receptors across the nuclear envelope.** *Proc Natl Acad Sci USA* 1993, **90:**3588-3592.
12. Baumann CT, Maruvada P, Hager GL, Yen PM: **Nuclear cytoplasmic shuttling by thyroid hormone receptors. Multiple protein interactions are required for nuclear retention.** *J Biol Chem* 2001, **276:**11237-11245.
13. Bunn CF, Neidig JA, Freidinger KE, Stankiewicz TA, Weaver BS, McGrew J, *et al.*: **Nucleocytoplasmic shuttling of the thyroid hormone receptor α.** *Mol Endocrinol* 2001, **15:**512-533.
14. Komeili A, O'Shea EK: **Nuclear transport and transcription.** *Curr Opin Cell Biol* 2000, **12:**355-360.
15. Tyagi RK, Amazit L, Lescop P, Milgrom E, Guiochon-Mantel A: **Mechanisms of progesterone receptor export from nuclei: role of nuclear localization signal, nuclear export signal, and Ran guanosine triphosphate.** *Mol Endocriol* 1998, **12:**1684-1695.
16. Liu J, DeFranco DB: **Protracted nuclear export of glucocorticoid receptor limits its turnover and does not require the exportin 1/CRM1-directed nuclear export pathway.** *Mol Endocrinol* 2000, **14:**40-51.
17. Holaska JM, Black BE, Love DC, Hanover JA, Leszyk J, Paschal BM: **Calreticulin is a receptor for nuclear export.** *J Cell Biol* 2001, **152:**127-140.
18. Burns K, Duggan B, Atkinson EA, Famulski KS, Nemer M, Bleackley RC, *et al.*: **Modulation of gene expression by calreticulin binding to the glucocorticoid receptor.** *Nature* 1994, **367:**476-480.
19. Dedhar S, Rennie PS, Shago M, Hagesteijn CY, Yang H, Filmus J, *et al.*: **Inhibition of nuclear hormone receptor activity by calreticulin.** *Nature* 1994, **367:**480-483.
20. Wheeler DG, Horsford J, Michalak M, White JH, Hendy GN: **Calreticulin inhibits vitamin D3 signal transduction.** *Nucleic Acids Res* 1995, **23:**3268-3274.
21. Omichinski JG, Clore GM, Schaad O, Felsenfeld G, Trainor C, Appella E, *et al.*: **NMR structure of a specific DNA complex of Zn-containing DNA binding domain of GATA-1.** *Science* 1993, **261:**438-446.
22. Katagiri Y, Takeda K, Yu ZX, Ferrans VJ, Ozato K, Guroff G: **Modulation of retinoid signalling through NGF-induced nuclear export of NGFI-B.** *Nat Cell Biol* 2000, **2:**435-440.
23. Hsieh JC, Shimizu Y, Minoshima S, Shimizu N, Haussler CA, Jurutka PW, *et al.*: **Novel nuclear localization signal between the two DNA-binding zinc fingers in the human vitamin D receptor.** *J Cell Biochem* 1998, **70:**94-109.
24. Freedman LP, Towers TL: **DNA binding properties of the vitamin $D_3$ receptor zinc finger region.** *Mol Endocinol* 1991, **5:**1815-1826.
25. McNally JG, Muller WG, Walker D, Wolford R, Hager GL: **The glucocorticoid receptor: rapid exchange with regulatory sites in living cells.** *Science* 2000, **287:**1262-1265.
26. DeFranco DB, Qi M, Borror KC, Garabedian MJ, Brautigan DL: **Protein phosphatase types 1 and/or 2A regulate nucleocytoplasmic shuttling of glucocorticoid receptors.** *Mol Endocrinol* 1991, **5:**1215-1228.
27. Doucas V, Shi Y, Miyamoto S, West A, Verma I, Evans RM: **Cytoplasmic catalytic subunit of protein kinase A mediates cross-repression by NF-κB and the glucocorticoid receptor.** *Proc Natl Acad Sci USA* 2000, **97:**11893-11898.
28. Migliaccio A, Castoria G, Di Domenico M, de Falco A, Bilancio A, Lombardi M, *et al.*: **Steroid-induced androgen receptor-oestradiol receptor β-SRC complex triggers prostate cancer cell proliferation.** *EMBO J* 2000, **19:**5406-5417.
29. Simoncini T, Hafezi-Moghadam A, Brazil DP, Ley K, Chin WW, Liao JK: **Interaction of oestrogen receptor with the regulatory subunit of phosphatidylinositol-3-OH kinase.** *Nature* 2001, **407:**538-541.
30. Kousteni S, Bellido T, Plotkin LI, O'Brien CA, Bodenner DL, Han L, *et al.*: **Nongenotropic, sex-nonspecific signaling through the estrogen or androgen receptors: dissociation from transcriptional activity.** *Cell* 2001, **104:**719-730.
31. Silver PA, Keegan LP, Ptashne M: **Amino terminus of the yeast GAL4 gene product is sufficient for nuclear localization.** *Proc Natl Acad Sci USA* 1984, **81:**5951-5955.
32. McBride KM, McDonald C, Reich NC: **Nuclear export signal located within the DNA-binding domain of the STAT1 transcription factor.** *EMBO J* 2000, **19:**6196-6206.
33. Carey KL, Richards SA, Lounsbury KM, Macara IG: **Evidence using a green fluorescent protein-glucocorticoid receptor chimera that the Ran/TC4 GTPase mediates an essential function independent of nuclear protein import.** *J Cell Biol* 1996, **133:**985-996.
34. Love DC, Sweitzer TD, Hanover JA: **Reconstitution of HIV-1 Rev nuclear export-independent requirements for nuclear import and export.** *Proc Natl Acad Sci USA* 1998, **95:**10608-10613.
35. Walker D, Htun H, Hager GL: **Using inducible vectors to study intracellular trafficking of GFP-tagged steroid/nuclear receptors in living cells.** *Methods* 1999, **19:**386-393.
36. Englmeier L, Olivo JC, Mattaj IW: **Receptor-mediated substrate translocation through the nuclear pore complex without nucleotide triphosphate hydrolysis.** *Curr Biol* 1999, **9:**30-41.
37. Kehlenbach RH, Dickmanns A, Gerace L: **Nucleocytoplasmic shuttling factors including Ran and CRM1 mediate nuclear export of NFAT In vitro.** *J Cell Biol* 1998, **141:**863-874.
38. Black BE, Lévesque L, Holaska JM, Wood TC, Paschal BM: **Identification of an NTF2-related factor that binds Ran-GTP and regulates nuclear protein export.** *Mol Cell Biol* 1999, **19:**8616-8624.
39. Black BE, Holaska JM, Lévesque L, Ossareh-Nazari B, Gwizdek C, Dargemont C, *et al.*: **NXT1 is necessary for the terminal step of Crm1-mediated nuclear export.** *J Cell Biol* 2001, **152:**141-155.

MOLECULAR AND CELLULAR BIOLOGY, Sept. 2002, p. 6286–6297
0270-7306/02/$04.00+0 DOI: 10.1128/MCB.22.17.6286–6297.2002

Vol. 22, No. 17

# $Ca^{2+}$-Dependent Nuclear Export Mediated by Calreticulin

James M. Holaska,[1,2]† Ben E. Black,[1,3]‡ Fraydoon Rastinejad,[3,4] and Bryce M. Paschal[1,3]*

*Center for Cell Signaling[1] and Departments of Microbiology,[2] Biochemistry and Molecular Genetics,[3] and Pharmacology,[4] University of Virginia, Charlottesville, Virginia 22908*

Received 22 January 2002/Returned for modification 6 March 2002/Accepted 23 May 2002

**We have characterized a pathway for nuclear export of the glucocorticoid receptor (GR) in mammalian cells. This pathway involves the $Ca^{2+}$-binding protein calreticulin (CRT), which directly contacts the DNA binding domain (DBD) of GR and facilitates its delivery from the nucleus to the cytoplasm. In the present study, we investigated the role of $Ca^{2+}$ in CRT-dependent export of GR. We found that removal of $Ca^{2+}$ from CRT inhibits its capacity to stimulate the nuclear export of GR in digitonin-permeabilized cells and that the inhibition is due to the failure of $Ca^{2+}$-free CRT to bind the DBD. These effects are reversible, since DBD binding and nuclear export can be restored by $Ca^{2+}$ addition. Depletion of intracellular $Ca^{2+}$ inhibits GR export in intact cells under conditions that do not inhibit other nuclear transport pathways, suggesting that there is a $Ca^{2+}$ requirement for GR export in vivo. We also found that the Ran GTPase is not required for GR export. These data show that the nuclear export pathway used by steroid hormone receptors such as GR is distinct from the Crm1 pathway. We suggest that signaling events that increase $Ca^{2+}$ could positively regulate CRT and inhibit GR function through nuclear export.**

The nuclear transport machinery integrates a variety of nuclear and cytoplasmic activities by mediating the translocation of housekeeping and regulatory proteins and RNAs. Translocation of these macromolecules generally requires a *cis*-acting transport signal, recognition of the signal by a receptor, and movement of the signal-receptor complex through the nuclear pore complex (NPC) (40, 45). In the case of nuclear export, most proteins rely on a hydrophobic nuclear export signal (NES) and its recognition by the receptor Crm1 (1, 10, 11, 39, 44). NES binding to Crm1 is stabilized by the presence of RanGTP, and the resulting trimeric complex of Crm1, NES, and RanGTP undergoes translocation through the nuclear pore (14, 28). Other proteins are important cofactors for this export pathway, including RanGAP, RanBP1, NXT1, and RanBP3 (1, 4, 22, 24).

A number of proteins that lack a hydrophobic NES are known to undergo nuclear export, and current evidence indicates that three distinct mechanisms can account for nuclear export of these proteins. First, an NES-containing adapter could be used to bridge the interaction between the protein and Crm1 (21). Second, the protein could use a different signal for nuclear export and undergo Crm1-independent export (3). Third, the protein could interact directly with nucleoporins in the NPC and undergo receptor-independent nuclear export (46). An advantage of these mechanisms is that they provide the potential for additional levels of regulation for protein sorting between the nucleus and cytoplasm.

* Corresponding author. Mailing address: Box 800577 Hospital West, Health Sciences System, University of Virginia, Charlottesville, VA 22908. Phone: (434) 243-6521. Fax: (434) 924-1236. E-mail: paschal@virginia.edu.

† Present address: Department of Cell Biology, Johns Hopkins University School of Medicine, Baltimore, MD 21205.

‡ Present address: Ludwig Institute for Cancer Research, University of California, San Diego, CA 92093.

The glucocorticoid receptor (GR) is an example of a protein that undergoes export from the nucleus even though it lacks a hydrophobic NES. Moreover, GR export is insensitive to the Crm1 inhibitor leptomycin B, which seems to rule out the use of NES adapters and Crm1 as the major receptor for this pathway (18, 25). The signal that specifies nuclear export of GR maps to the 67-amino-acid DNA binding domain (DBD), which is both sufficient to mediate the export of green fluorescent protein (GFP) reporter proteins and necessary for export of GR (3). Mutations that disrupt DBD structure, whether in the context of an isolated DBD or in full-length GR, also reduce its export activity. The most severe mutations that reduce DBD-dependent nuclear export are two phenylalanine-to-alanine mutations in the DNA recognition helix (3). The structural conservation of the DBD in the nuclear receptor superfamily suggests that the DBD could be widely used as an export signal. Support for this hypothesis was obtained by showing that the DBDs from steroid, nonsteroid, and orphan nuclear receptors can function as export signals when fused to a GFP reporter protein (3).

Nuclear export mediated by the DBD of GR involves a $Ca^{2+}$ binding protein named calreticulin (CRT), which was first described as a protein in the lumen of the endoplasmic reticulim (ER) (29). The original link between these proteins came with the finding that a peptide sequence, KLGFFKR, which is recognized by CRT, is related to the peptide sequence KVFFKR, which is found in the DNA recognition helix of GR (33). CRT binds to GR and blocks its interaction with DNA in gel shift experiments, and overexpression of CRT inhibits GR-dependent transcription in cells (6). Similar results were obtained when the interactions between the androgen, retinoic acid, and vitamin D receptors and CRT were examined (8). The initial interpretation of these results was that a pool of CRT outside of the lumen of the ER acts as a negative regulator of transcription. The technical difficulty of showing that CRT is out-

side the ER, however, gave way to the view that CRT is restricted to the ER lumen and that the effects of CRT on gene expression are indirect (23).

Our laboratory purified CRT from HeLa cell cytosol in a search for novel export factors (18). We used a permeabilized cell assay that measures the nuclear export of protein kinase inhibitor (PKI), a protein that contains a hydrophobic NES (19). CRT can bind directly to the NES in PKI, which contains the peptide sequence LALKLAGLDIN. In binding experiments the interaction of CRT with PKI is stabilized in the presence of RanGTP, and in permeabilized cells CRT-dependent export of PKI is enhanced by RanGTP (18). The ability of CRT to act as an export factor for PKI and its prior link with steroid receptor function led us to test whether CRT can function as an export factor for GR. We found that recombinant CRT can stimulate GR export and that the GR export deficiency in *crt*$^{-/-}$ cells can be rescued with CRT (18). These observations support the view that CRT functions as an export factor for GR.

In the present study we characterized the CRT-dependent nuclear export of GR, with particular emphasis on how this pathway could be regulated. Because CRT is a $Ca^{2+}$ binding protein, we have examined how removing $Ca^{2+}$, either by chelation or by deletion of $Ca^{2+}$ binding domains, affects CRT binding to the GR DBD and CRT-dependent nuclear export of GR. We show that $Ca^{2+}$ binding to CRT is necessary for direct binding to the DBD and for nuclear export of GR in permeabilized cells. This $Ca^{2+}$ requirement involves the C-terminal domain of CRT, which contains multiple low-affinity, high-capacity binding sites for $Ca^{2+}$ (2). Removal of the C-terminal domain renders CRT insensitive to $Ca^{2+}$ chelation, suggesting that this domain performs a regulatory function that is linked to $Ca^{2+}$ binding. While the $Ca^{2+}$-loaded CRT is active for GR export and $Ca^{2+}$-free CRT is inactive for GR export, the opposite is the case for NES export in permeabilized cells. Thus, $Ca^{2+}$-loaded CRT is inactive for Rev export and $Ca^{2+}$-free CRT is active for Rev export. The $Ca^{2+}$-loaded and $Ca^{2+}$-free forms of CRT were shown to have different sensitivities to proteases (7), indicating that CRT adopts different protein conformations. Our data show that these $Ca^{2+}$-dependent conformations of CRT are correlated with the ability to recognize different protein substrates. $Ca^{2+}$ depletion in vivo results in the inhibition of GR export, consistent with a $Ca^{2+}$ requirement for the translocation of GR from the nucleus to the cytoplasm. Our finding that CRT requires $Ca^{2+}$ for binding and nuclear export of GR suggests a potential mechanism for regulating this pathway.

## MATERIALS AND METHODS

**Plasmids and recombinant proteins.** Standard methods were used for the expression and purification of glutathione *S*-transferase (GST)- and His-tagged proteins, all of which were stored at −80°C as single-use aliquots. The plasmid encoding the GST fusion with CRT was constructed in pGEX4T3, using the open reading frame of mouse CRT lacking the N-terminal 17-amino-acid signal sequence. GST-CRT was expressed in DH5α bacteria as described previously (18). Full-length and deletion mutants of mouse CRT were cloned into pQE30, and the His-tagged proteins were expressed in TG1 bacteria (3). The wild-type (WT) and NES mutant forms of PKI (L41A, L44A) (44) were expressed in BL21(DE3) bacteria and purified without fusion tags as described previously (19). The plasmid encoding a hydrophobic NES was generated by cloning the DNA sequence that encodes residues 35 to 49 of human PKI into pGEX4T3. Likewise, the plasmid encoding a nuclear receptor DBD was generated by cloning the DNA sequence that encodes residues 413 to 509 of human GR (3). The GST-NES and GST-DBD proteins immobilized on glutathione beads were either used directly for binding experiments or eluted and used in microtiter plate binding assays. For certain competition experiments, WT and mutant (MUT) forms of the GR DBD (FF to AA) were expressed as GST fusion proteins, cleaved from GST using thrombin, and further purified by ion-exchange chromatography (3). Plasmids encoding His-tagged Ran (WT, Q69L, and T24N; a gift of D. Görlich) and His-tagged Crm1 (a gift of L. Gerace) were used to express proteins in TG1 bacteria, which were purified on Talon resin (Clontech).

**Binding assays.** The three formats used for the solid-phase binding assays were microtiter wells, biosensor cuvettes, and Sepharose beads. The microtiter well assay was performed as described previously (4). Briefly, purified target proteins were immobilized in high-binding 96-well plates (Costar no. 3590) overnight at 4°C in 1× transport buffer (30). Unbound protein was removed, and the plates were blocked overnight with bovine serum albumin (30 mg/ml). Binding assays (with 100-μl mixtures) were performed in triplicate, using radiolabeled Ran and CRT or Crm1, and the level of binding was measured by scintillation counting as described previously (4). The biosensor assay was performed as described previously (18). Briefly, biotinylated NES peptide was immobilized in streptavidin-coated cuvettes (Fisons) for 15 min at room temperature. The cuvettes were washed with PBS and used for binding assays with the proteins indicated in the legend to Fig. 3. A detailed description of the Fisons biosensor, which measures the change in refractive index that occurs on protein-ligand binding and dissociation, has been published (34). The Sepharose bead binding assays, using either glutathione beads and GST proteins or Talon beads and His-tagged proteins, were carried out by standard methods. Briefly, proteins were immobilized on the beads and blocked overnight with bovine serum albumin (30 mg/ml), and the assays (with 100-μl mixtures) were performed using the proteins indicated in the legends. The bound fractions were examined by immunoblotting using antibodies to CRT or Crm1 and enhanced chemiluminescence.

**$Ca^{2+}$ removal from CRT.** $Ca^{2+}$ was removed from CRT by a published procedure that involves treatment with EGTA (43). Recombinant CRT (0.5 mg/ml in PBS) was incubated for 10 min at 30°C in the presence of 10 mM EGTA, and the sample was then transferred to ice. $Ca^{2+}$ was rebound to CRT by supplementing half of the EGTA-treated sample with excess $CaCl_2$ (final $Ca^{2+}$ concentration, 20 mM). The EGTA- and $Ca^{2+}$-treated samples were used at a dilution of at least 1:10 in nuclear export and binding assays, such that the maximum concentration of EGTA in the assays was ~1 mM.

**GR export assays.** Nuclear export of GR in permeabilized cells was performed essentially as described previously (18), except that a stable cell line expressing GR-GFP (a gift of G. Hager) was used instead of transiently transfected cells. The cell line (3676 cells) (27) expresses GR-GFP under the control of a tetracycline-regulated promoter. The cells were grown on glass coverslips for 16 h in the absence of tetracycline to allow GR-GFP expression, and nuclear import of GR-GFP was induced in vivo by dexamethasone (Dex; 1μM) addition to the media. The cells were permeabilized with digitonin (0.005%) for 5 min and used for export assays in vitro with the combinations of transport factors indicated in the legends. At the end of the 20-min export reaction, the samples were washed, fixed, stained with 4′,6-diamidino-2-phenylindole (DAPI), and mounted on glass slides. Using a Nikon Microphot SA microscope (60× objective, numerical aperture N.A. = 1.40) and a Hamamatsu C-4742-95 charge-coupled device camera, ~50 nuclei from each coverslip were selected using the DAPI channel and DAPI and GFP images were captured. The digital images were acquired using Openlab 2.06 on a Macintosh G3 computer (OS 9.0), and figures were assembled using Adobe Photoshop 5.5 and Freehand 9.0. Mixtures for reactions that measured export and import in the same nuclei contained, in addition to CRT or Crm1, the fluorescent protein allophycocyanin coupled with NLS peptide (APC-NLS) and recombinant import factors (importins and NTF2).

**Hydrophobic NES export assays.** The assay for nuclear export of Rev in permeabilized cells has been described previously (RGG2.2 cells) (18, 26). The fluorescent reporter in this cell line (denoted Rev-GFP) also contains the ligand binding domain of GR, which confers Dex-inducible import in vivo. Digitonin permeabilization, nuclear export, and analysis using fluorescence microscopy were performed as described above for GR-GFP. All export assay mixtures contained Ran, NXT1, and RanBP1. In experiments designed to test the role of Ran in CRT-dependent export, the T24N and Q69L mutant forms of Ran were preloaded with unlabeled GDP and GTP, respectively, and unincorporated nucleotide was removed on a desalting column.

**$Ca^{2+}$ depletion in living cells.** Nuclear export of GR was assayed under conditions of $Ca^{2+}$ depletion in the 3676 cells by measuring the net redistribution from the nucleus to the cytoplasm. Cells expressing GR-GFP were treated with Dex for 1 h to induce nuclear import and subsequently transferred to phenol red-free media containing ionomycin (Ion; 1 μM) or thapsigargin (TG; 1 μM) in

the presence of 1,2-bis(*O*-aminophenoxy)ethane-*N*,*N*,*N*', N'-tetraacetic acid-acetoxymethyl ester (BAPTA-AM; 10 μM). Cells were also treated with dimethylsulfoxide (DMSO; 0.1%) as a control since the Ion and TG were prepared as 1,000× stocks in DMSO. At 0-, 3-, and 6-h time points, the coverslips were fixed in formaldehyde (3.7%) and mounted using Vectashield. The ratio of nuclear to cytoplasmic GR-GFP fluorescence was measured using Openlab software in at least 50 cells per condition. Similar methods were used to assay the effect of $Ca^{2+}$ depletion on Rev-GFP export using the RGG2.2 cells.

## RESULTS

**Assembly of CRT complexes in vitro.** In digitonin-permeabilized cells, CRT can stimulate the nuclear export of proteins such as PKI that contain a hydrophobic NES, in a reaction that is dependent on Ran (18). In addition, CRT can stimulate the nuclear export of steroid hormone receptors that lack a hydrophobic NES. The export signal for CRT-dependent export of steroid hormone receptors is contained within the DBD of these proteins (3). The lack of apparent structural similarity between the hydrophobic NES and the DBD led us to characterize how CRT can recognize distinct export signals.

Because Ran is a stoichiometric component of export complexes that contain Crm1 and NES proteins, we first examined whether Ran might also function as a component of export complexes that contain CRT. Previously, we have shown that Ran assembles into a complex containing NES and CRT (18). In the present study, we tested whether Ran can assemble into a complex containing the DBD of GR (amino acids 413 to 509) and CRT. WT PKI, NES MUT PKI, and GST-DBD were immobilized in microtiter wells and incubated with recombinant CRT or Crm1 (5 μg each) in the presence of radiolabeled RanGTP ($2 \times 10^4$ cpm). RanGTP assembled into a complex with CRT or Crm1 in the presence of WT PKI. The reaction is specific since a functional NES in PKI is required for complex assembly. In contrast to these results obtained with the NES-containing protein PKI, RanGTP did not coassemble into a complex with CRT or Crm1 in the presence of the DBD from GR (Fig. 1A). Thus, CRT binding to the DBD is qualitatively different from CRT binding to the hydrophobic NES, since only the latter involves RanGTP as a stoichiometric component.

We carried out binding reactions with the DBD immobilized on glutathione beads and confirmed that CRT binds directly to the DBD and that binding is neither enhanced nor prevented by RanGTP (Fig. 1B). Under the reaction conditions, Crm1 binding to the hydrophobic NES was stimulated by the presence of RanGTP (lane 6). We did not observe Crm1 binding to the DBD in the absence or presence of RanGTP, consistent with our data that Crm1 is not the receptor for proteins that use the DBD as an export signal (18). Our data demonstrate that CRT can assemble into two types of export complexes in vitro: a CRT-NES-RanGTP complex and a CRT-DBD complex.

**Ran-dependent and Ran-independent export by CRT.** The results of our binding assays suggested that CRT-mediated export might occur by both Ran-independent and Ran-dependent mechanisms. That is, the data suggested that nuclear export of hydrophobic NES-containing proteins could be Ran dependent and nuclear export of DBD-containing proteins could be Ran independent. We tested this hypothesis by manipulating the composition and concentration of recombinant export factors in permeabilized cell assays, using GFP fusions of Rev and GR as the export substrates to assay NES- and

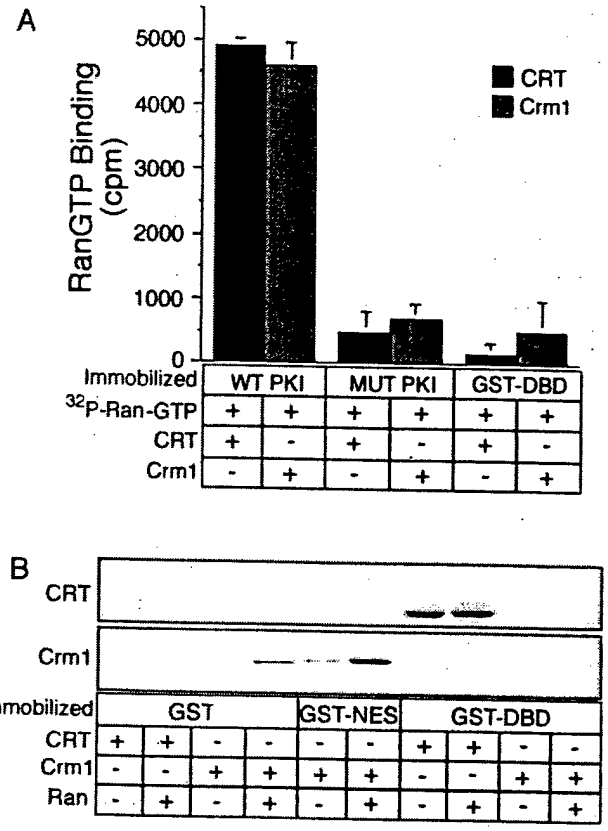



FIG. 1. Formation of export complexes involving CRT and Crm1. (A) The incorporation of RanGTP into complexes containing CRT and Crm1 was assayed using Ran preloaded with [γ-$^{32}$P] GTP. Target proteins were immobilized in microtiter wells (500 ng/well), and CRT or Crm1 (5 μg each) was added to each well together with $2 \times 10^4$ cpm of radiolabeled RanGTP. Following incubation for 1 h at room temperature, the wells were washed four times and the bound fractions were released and assayed by scintillation counting. (B) RanGTP is not a cofactor for CRT binding to the DBD. Target proteins (1 μg each) were immobilized on glutathione beads, and CRT or Crm1 (500 ng each) was added to each sample in the absence or presence of Ran (1 μg) preloaded with cold GTP. The samples were mixed end over end for 2 h at room temperature, washed three times, eluted, and analyzed by immunoblotting with antibodies to CRT and Crm1. These data show that, like Crm1, NES recognition by CRT involves RanGTP. In contrast, DBD recognition by CRT does not involve RanGTP.

DBD-dependent export, respectively. Both GFP export substrates contain the hormone binding domain of GR, which directs efficient Dex-dependent nuclear import in vivo (31). Following a digitonin permeabilization step, nuclear export of the GFP reporters can be stimulated by the addition of soluble transport factors (26).

We tested for the role of Ran in CRT-dependent NES export by using Rev, which contains a hydrophobic NES (LPPLERLTL) (10). We used a concentration of CRT that was determined to be subsaturating with respect to Rev-GFP export. In the presence of 0.2 μM CRT, WT Ran preloaded with GTP stimulated Rev-GFP export (Fig. 2A). In contrast, the Ran mutant T24N, which mimics the GDP-bound form of Ran, was inactive for CRT-dependent export. The Ran mutant Q69L, which mimics the GTP-bound form of Ran, showed little effect on export mediated by CRT, which is different from

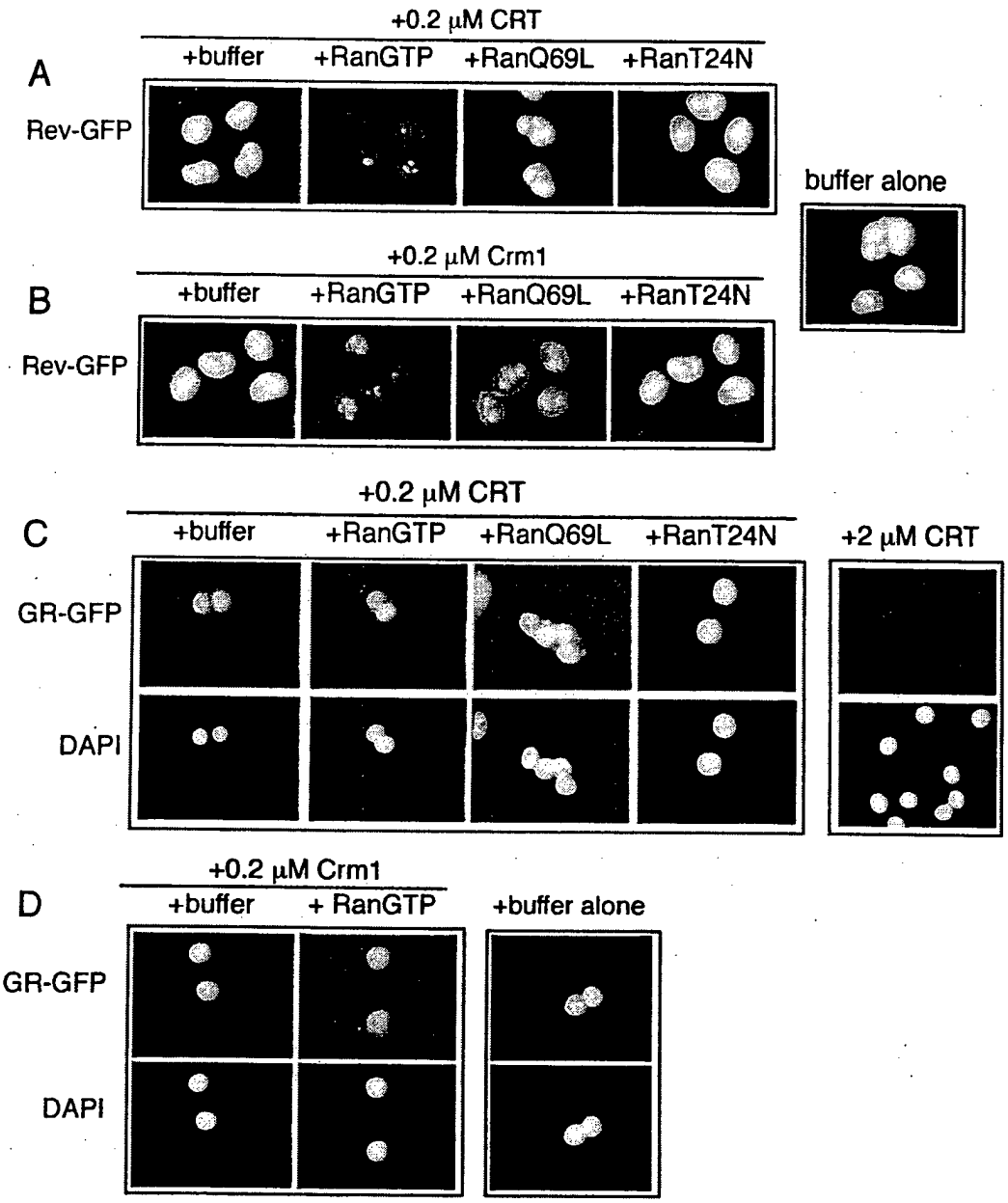



FIG. 2. Ran is a cofactor for NES-dependent export but not for DBD-dependent export. Nuclear export was assayed by supplementing digitonin-permeabilized cell assay mixtures with CRT (0.2 or 2.0 μM), Crm1 (0.2 μM), and different forms of Ran (1.9 μM each). NES-dependent export was assayed using a cell line that expresses a GFP fusion with Rev (RGG2.2) (26), and DBD-dependent export was assayed using a cell line that expresses a GFP fusion with GR (3676) (27). The Rev-GFP fusion also contains the ligand binding domain of GR, which imparts ligand-dependent nuclear import of the reporter protein. Prior to digitonin permeabilization, both cell lines were treated with Dex (1 μM) to induce nuclear import of the GFP reporters. (A and B) RanGTP stimulates Rev-GFP export in the presence of CRT or Crm1. (C and D) In contrast, neither WT nor MUT forms of Ran stimulate GR-GFP export in the presence of CRT or Crm1.

previous results obtained using another RanGTP mutant (G19V) and a fluorescent conjugate of PKI as the export substrate (18). The reason for this discrepancy is unclear, but it could relate to subtle differences in the structure and activity of these mutant proteins or to the fact that different NES proteins were used in the two assays, or both. Crm1 in these assays (Fig. 2B) supports robust export of Rev-GFP in the presence of WT Ran preloaded with GTP. In the presence of the Ran mutant Q69L, a moderate level of export was observed, while the Ran mutant T24N failed to support Crm1-dependent export. Thus, Ran stimulates both CRT- and Crm1-dependent export in this system, and maximal export activity of either protein is observed only in the presence of WT Ran.

GR-GFP export was assayed at the same concentration of CRT (0.2 μM) that, in the presence of RanGTP, promotes efficient export of Rev GFP. We determined that the addition of subsaturating CRT together with Ran does not support nuclear export of GR. We also found that the addition of Crm1

and RanGTP at concentrations (0.2 and 1.9 μM, respectively) that promote efficient Rev export fails to support GR export. These transport data correlate with our in vitro binding data, since CRT-dependent binding and export of an NES-containing protein is dependent on RanGTP. In contrast, CRT-dependent binding and export of the DBD appears to be independent of Ran. In the context of permeabilized cell assays, CRT can substitute for Crm1 in the NES pathway; however, Crm1 cannot substitute for CRT in the DBD pathway.

**Overlapping export substrate binding sites on CRT.** We considered two models to account for the ability of CRT to bind and mediate the nuclear export of two substrates with apparently unrelated export signals. In the first model, CRT could contain two distinct substrate binding sites, with only one of the two sites regulated by RanGTP. In the second model, CRT could contain a single substrate- binding site that accommodates both types of substrate and could use RanGTP to regulate the binding of only one of the two types of substrate. We reasoned that if CRT uses a single substrate binding site (or a single type of binding site), then CRT binding to the DBD should be inhibited in the presence of excess NES. To address this issue, we assayed CRT binding to an immobilized DBD in the absence and presence of WT and NES mutant forms of PKI. CRT binding to the DBD was reduced by the presence of WT PKI but not by the NES mutant form of PKI (Fig. 3A). Our data are consistent with a single type of substrate binding site on CRT, although we cannot rule out the possibility that CRT contains a second substrate binding site that is inhibited by an allosteric mechanism.

We also tested whether CRT contains a single type of substrate binding site by assaying nuclear export in permeabilized cells. Nuclear export of GR-GFP was tested in the absence and presence of WT and NES mutant forms of PKI. Inclusion of excess WT PKI (12 μM) in the reaction mixture inhibited the nuclear export of GR, whereas inclusion of the NES mutant form of PKI at the same concentration had no effect (Fig. 3B). GR export mediated by CRT was blocked by excess GR DBD (4.5 μM), consistent with previous data showing that the DBD functions as the signal for nuclear export of GR (3, 18).

The assays described above (Fig. 3A and B) were designed to measure the effect of excess NES on CRT binding to, and export of, the DBD. We performed analogous protein binding and nuclear export experiments that, instead, measured the effect of excess DBD on the interaction between CRT and the NES (Fig. 3C). In the biosensor assay, CRT binding to the NES (green tracing) or Crm1 binding to the NES (purple tracing) were both observed, but only in the presence of Ran GTP. CRT binding to the NES was reduced markedly by the presence of a WT DBD (light blue tracing), but this interaction was unaffected by the presence of a MUT DBD (dark blue tracing). The ability of the DBD to compete with NES for binding to CRT provides evidence that a single type of binding site is used for both substrates.

The effect of excess DBD on CRT-dependent export of NES substrate was tested in permeabilized cells by using Rev-GFP as a reporter. CRT-dependent export of Rev-GFP was inhibited by excess DBD and, as expected, by excess WT PKI (Fig. 3D). Addition of excess WT PKI resulted in inhibition of Crm1-dependent export of Rev-GFP, while addition of excess DBD had no effect on Rev-GFP export (data not shown). In summary, our results demonstrate that (i) CRT can bind to and mediate the nuclear export of a protein that contains either a hydrophobic NES or a steroid hormone receptor DBD, (ii) CRT-dependent binding and export of a protein containing the hydrophobic NES is dependent on Ran, (iii) CRT-dependent binding and export of a protein containing an appropriate DBD is not dependent on Ran, and (iv) CRT uses a similar substrate binding site for proteins that contain either a hydrophobic NES or an appropriate DBD.

**$Ca^{2+}$ is critical for CRT export activity.** CRT was originally discovered as a $Ca^{2+}$ binding protein (29), and many of its functions in the ER have been proposed to be linked to its ability to bind $Ca^{2+}$ (20). To determine if $Ca^{2+}$ influences the nuclear export activity of CRT, we used an established method to release $Ca^{2+}$ from recombinant CRT (43) and tested the protein in binding and transport assays. EGTA-induced $Ca^{2+}$ release from CRT resulted in the loss of CRT binding to the DBD. Binding to the DBD was, however, restored by the addition of excess $Ca^{2+}$ (Fig. 4A). $Ca^{2+}$ removal from CRT caused a corresponding reduction in its capacity to promote nuclear export of GR in permeabilized cells, and the export activity was restored by the addition of excess $Ca^{2+}$ (Fig. 4B). Unexpectedly, the capacity of CRT to mediate nuclear export of the NES substrate Rev-GFP was unaffected by EGTA, and the addition of excess $Ca^{2+}$ to this assay actually inhibited Rev-GFP export (Fig. 4C). The effect of $Ca^{2+}$ on Rev-GFP export appears to be linked to CRT since these treatments did not affect Crm1-dependent export of Rev-GFP (Fig. 4D). These data suggest that the $Ca^{2+}$-bound state of CRT determines whether it binds to DBD- or NES-containing proteins and that these interactions are mutually exclusive under the conditions used in our assays.

We considered it formally possible that addition of EGTA to digitonin-permeabilized cells could block GR export by inhibiting NPC function. $Ca^{2+}$ depletion in cells has been reported to inhibit the nuclear import of an NLS reporter protein (15), and there is structural evidence that depletion of $Ca^{2+}$ from the lumen of the nuclear envelope can alter NPC structure (41). It was also possible that $Ca^{2+}$-free CRT might accumulate at the NPC and block transit through the nuclear pore. We addressed this issue in a permeabilized cell assay that reconstitutes both nuclear import and export (Fig. 5A), reasoning that nuclear import would provide a relevant readout of NPC activity that is independent of nuclear export. Import in these assays was reconstituted by the addition of α- and β-importin, Ran, and NTF2 and was monitored by the nuclear accumulation of allophycocyanin conjugated with the simian virus 40 large T-antigen NLS (APC-NLS). EGTA effectively blocked GR export without affecting APC-NLS import (Fig. 5B). Moreover, excess $Ca^{2+}$ blocked CRT-dependent export of Rev-GFP without affecting APC-NLS import. This effect was specific to CRT since Crm1-dependent export of Rev-GFP was unaffected by excess $Ca^{2+}$ (Fig. 5C). Our data are consistent with $Ca^{2+}$-mediated regulation of the activity of CRT and not with the inactivation of NPC function. We note that our standard assay buffer contains 1 mM EGTA. Because this condition is permissive for CRT-dependent export, removal of $Ca^{2+}$ from CRT may require the higher concentration of EGTA (10 mM) or pretreatment of purified CRT with EGTA at 30°C, or both.

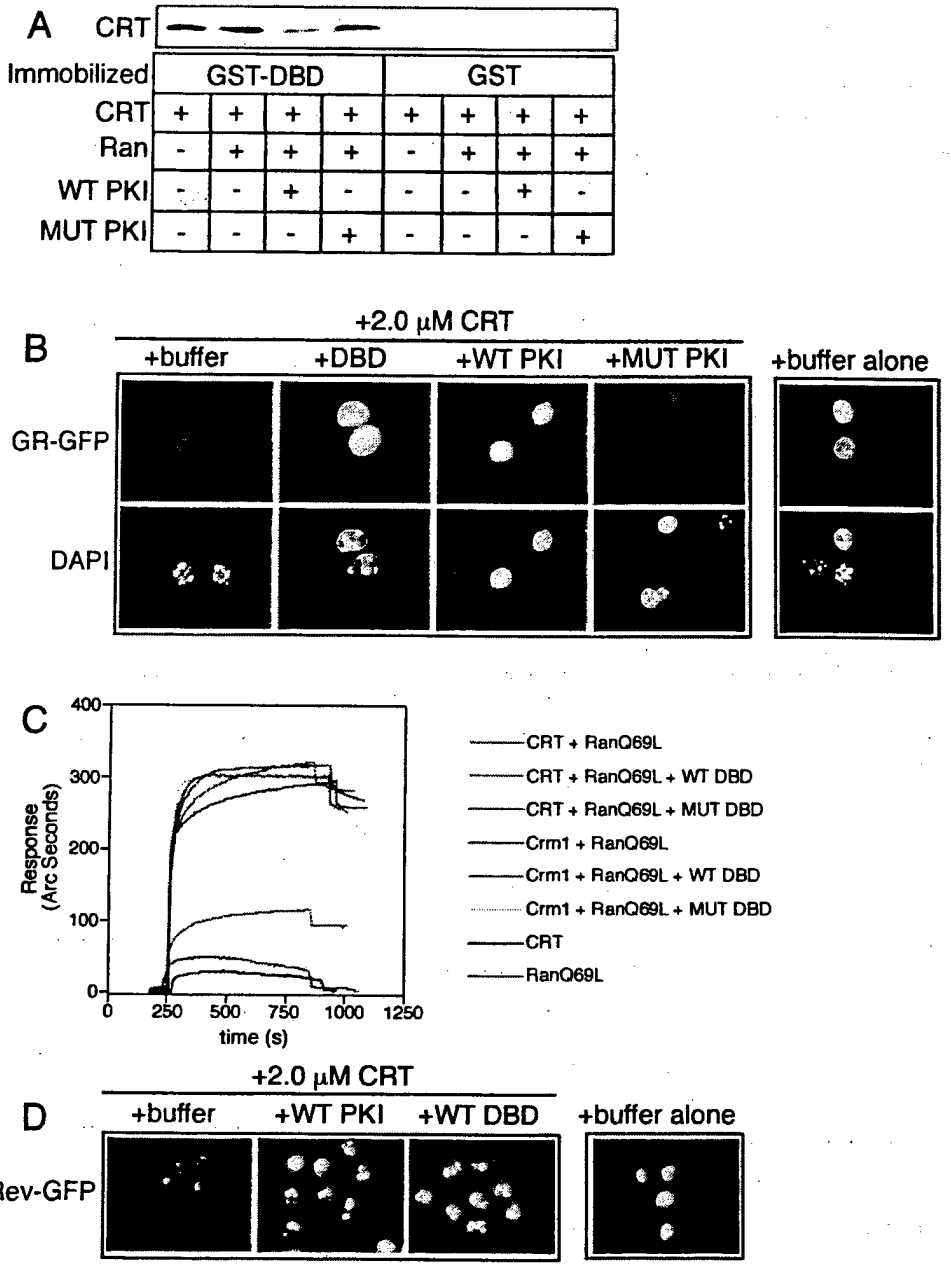



FIG. 3. The DBD and hydrophobic NES use a common or overlapping binding site on CRT. (A) Binding assay with GST-DBD or GST (2 μg each) immobilized on glutathione beads and CRT (500 ng), RanGTP (2 μg), and WT or MUT PKI (2 μg) added in solution. The bound fractions were analyzed by immunoblotting for CRT. Including WT PKI in the reaction reduced the level of CRT bound to the DBD, indicating that these proteins bind to similar sites on CRT. This competition was not observed when RanGTP was omitted from the assay (data not shown), consistent with Ran acting as a cofactor for NES binding but not for DBD binding. (B) Nuclear export of GR-GFP was assayed in permeabilized cells using CRT (2.0 μM) in the presence of buffer, excess DBD (4.5 μM), or PKI (12 μM WT or MUT). (C) Competitive binding interactions between NES, DBD, and CRT measured in a biosensor assay. NES peptide was immobilized on the cuvette surface through a biotin-neutravidin linkage and was used to measure the Ran-dependent binding of CRT in the absence and presence of DBD in the solution. The proteins used in the assay were CRT (1.1 μM), Crm1 (1.1 μM), RanQ69L (1.9 μM), and WT and MUT DBD from GR (9.1μM each). CRT binds efficiently to NES peptide in the presence of Ran (green tracing), and this can be competed with the WT DBD (light blue tracing) but not with the transport-defective MUT DBD that contains the FF-to-AA mutations (dark blue tracing). Crm1 binding to the NES is unaffected by the presence of excess DBD (fuchsia tracing). (D) The DBD competes with NES in the CRT-dependent export pathway. The cell line expressing Rev-GFP was used to assay export in the presence of CRT (1.1 μM), WT PKI (12 μM), and WT DBD (9.1 μM) as indicated.

**The P-domain and C-terminal domain of CRT Impart $Ca^{2+}$ regulation.** As diagrammed in Fig. 6A, the domain structure of CRT includes an acidic C-terminal domain that contains multiple low-affinity, high-capacity $Ca^{2+}$ binding sites ($K_D \approx 2$ mM; >25 mol/mol) and a proline-rich P-domain that contains high-affinity, low-capacity $Ca^{2+}$ binding sites ($K_D \approx 1$ μM) (2). In light of our data showing that $Ca^{2+}$ can regulate CRT-dependent export of GR, we examined whether the C-terminal

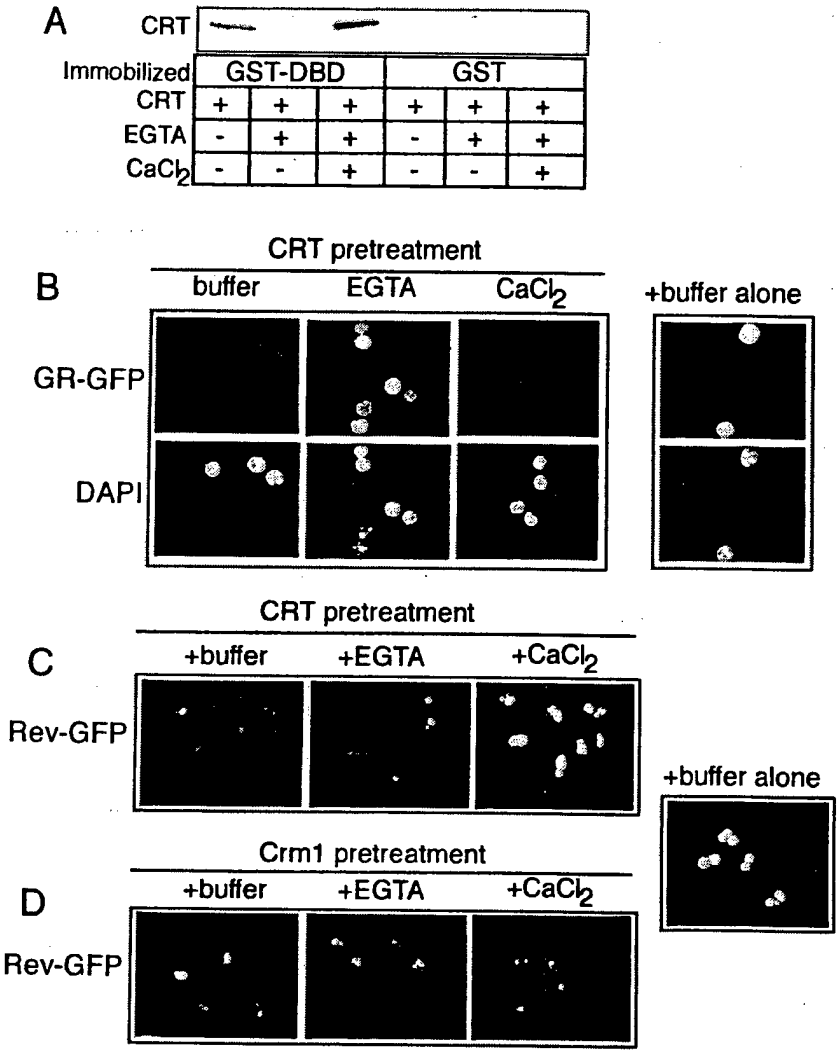



FIG. 4. $Ca^{2+}$ binding to CRT is necessary for nuclear export of GR. (A) Binding assay performed with GST-DBD or GST (2.5 μg each) immobilized on glutathione beads and CRT (500 ng). The amount of CRT bound in each reaction fraction was examined by immunoblotting for CRT. The CRT used in the binding assay was untreated, pretreated with 10 mM EGTA, or pretreated with 10 mM EGTA and 20 mM $CaCl_2$. $Ca^{2+}$ removal from CRT inhibits binding to the DBD, and this can be reversed by addition of $Ca^{2+}$. (B and C) $Ca^{2+}$ is required for CRT-dependent GR export; however, $Ca^{2+}$ inhibits CRT-dependent NES export. (D) The presence of excess EGTA and $Ca^{2+}$ in the permeabilized-cell assay mixture does not affect the export mediated by Crm1. CRT and Crm1 were used at a final concentration of 1.1 μM each in the export assays. The pretreatment of CRT and Crm1 with EGTA is described in Materials and Methods.

domain is required for CRT activity in our assays. CRT mutants lacking a portion of the C-terminal domain (retaining residues 1 to 350) or the entire C-terminal domain (retaining residues 1 to 273) were still functional for nuclear export of Rev-GFP, and the activities of these mutants was blocked by excess $Ca^{2+}$ (Fig. 6B to D). From these data we infer that the $Ca^{2+}$-dependent inhibition of hydrophobic NES export by CRT can be ascribed to the high-affinity low-capacity $Ca^{2+}$ binding site in the P-domain.

The deletion mutants were also examined for the $Ca^{2+}$ dependence of DBD binding and GR export. We found that the deletion mutants still bound to the DBD in the presence of $Ca^{2+}$ and that the mutant (residues 1 to 273) lacking the entire C-terminal domain was able to bind to the DBD in the presence of EGTA, albeit to a lesser extent than was the full-length CRT (Fig. 6E). This suggests that in the absence of $Ca^{2+}$, the C-terminal domain normally exerts a negative regulation on the substrate binding activity of CRT. Removal of the entire C-terminal domain (residues 1 to 273) also resulted in loss of the EGTA-dependent inhibition of GR export that is observed with both the full-length and partial C-terminal domain deletion mutant (Fig. 6F to H). These data indicate that the low-affinity, high-capacity $Ca^{2+}$ binding sites in the C-terminal domain of CRT are not required for nuclear export of hydrophobic NES or DBD-containing proteins but that they do contribute a regulatory function. The $Ca^{2+}$ binding sites in the C-terminal domain are important for the EGTA-sensitive GR export, and the $Ca^{2+}$ binding site in the P-domain appears to be sufficient to confer $Ca^{2+}$-induced inhibition of NES export.

**GR-GFP Export is Sensitive to $Ca^{2+}$ Depletion.** Our results showing that GR undergoes CRT- and $Ca^{2+}$-sensitive export in permeabilized cells prompted us to examine whether this

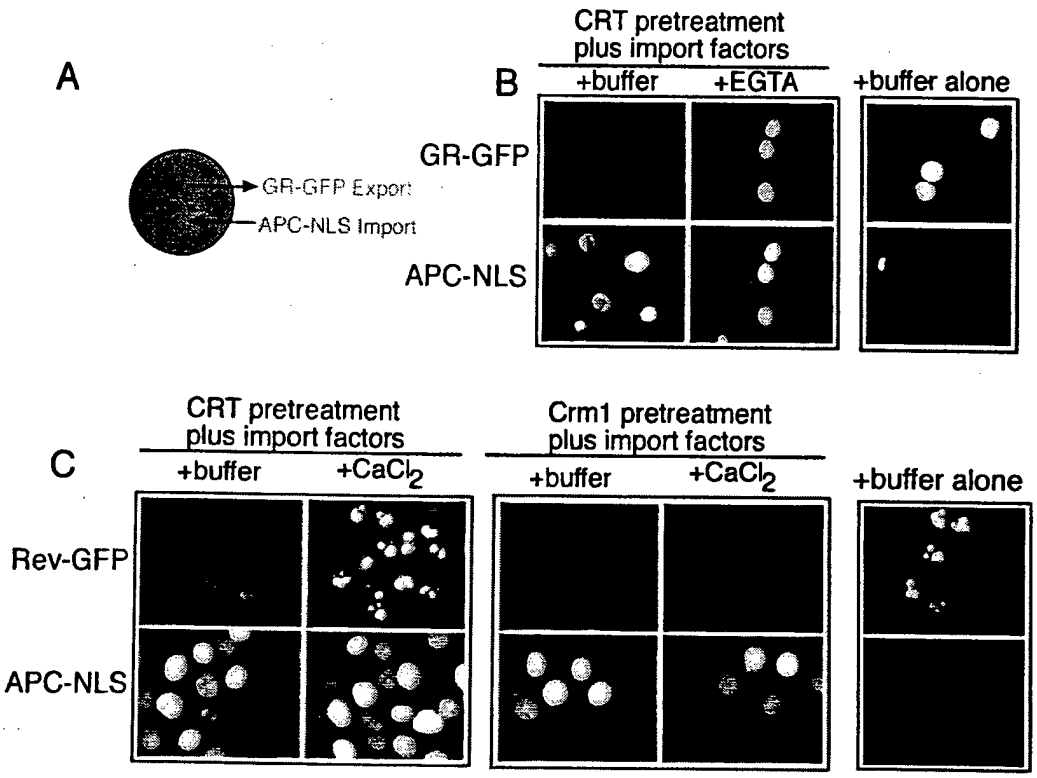



FIG. 5. $Ca^{2+}$ chelation with EGTA inhibits CRT-dependent export without affecting nuclear import in permeabilized cells. (A) Diagram illustrating the assay for nuclear import (APC-NLS) and export (GR-GFP) in the same cells. (B) EGTA blocks CRT-dependent export of GR-GFP without affecting the nuclear import of APC-NLS in the presence of β-importin, Rch1, Ran, NTF2, NXT1, and RanBP1 (50μg/ml each). (C) Excess $Ca^{2+}$ blocks CRT-dependent export of Rev-GFP without affecting the nuclear import of APC-NLS.

pathway is operational in living cells. The 3676 cell line (27), which stably expresses GR-GFP, was treated with Dex to induce import and then transferred to Dex-free medium that contained reagents known to deplete luminal and cytosolic $Ca^{2+}$. The effect of $Ca^{2+}$ depletion on nuclear export was examined by fluorescence microscopy after 6 h and quantitated by measuring the nuclear/cytoplasmic fluorescence ratio of GR-GFP. The conditions included treatment with the $Ca^{2+}$ ionophore Ion (1 μM) or the ER $Ca^{2+}$ pump inhibitor TG (1 μM). The membrane-permeable form of BAPTA-AM (10 μM) was included with Ion and TG to chelate $Ca^{2+}$ that was released. We found that Ion or TG administered in the presence of BAPTA-AM were both effective at blocking the nuclear export of GR-GFP (Fig. 7A). During the export phase of the experiment, the nuclear/cytoplasmic fluorescence in the control cells showed a significant reduction from 11.9 ± 2.1 to 1.5 ± 1.5 (Fig. 7B). In contrast, the nuclear/cytoplasmic fluorescence in the Ion- and TG-treated cells showed only a slight reduction, to 9.9 ± 1.2 and 10.5 ± 0.9, respectively.

It is unlikely that Ion, TG, and BAPTA-AM cause a global disruption of NPC structure because NES-dependent export of Rev-GFP continues under $Ca^{2+}$ depletion conditions in vivo (Fig. 7C) as well as in vitro (26). The transport of other cargos is unaffected by $Ca^{2+}$ depletion in vivo, including Crm1-dependent nuclear export of mitogen-activated protein kinase-activated protein kinase 2 (42) and importin/karyopherin-dependent nuclear import of Rev-GFP (B. Black, unpublished observations). Our finding that nuclear export of GR in permeabilized cells is sensitive to the $Ca^{2+}$-bound state of the CRT used in the assays suggests that depletion of $Ca^{2+}$ in vivo blocks the CRT-dependent export of GR in the cell.

## DISCUSSION

Our study has revealed that $Ca^{2+}$ and RanGTP can regulate the nuclear export activity of CRT. $Ca^{2+}$ binding to CRT is necessary for its interaction with the DBD of GR in solid-phase binding assays and for nuclear export of GR in digitonin-permeabilized cell assays. These interactions appear to be independent of Ran, since Ran is not required for CRT binding to the DBD. Ran does not stimulate GR export when CRT is rate-limiting in the assay, and GTPase-deficient mutants of Ran do not block GR export. Under conditions where the low-affinity $Ca^{2+}$ binding sites of CRT are saturated with $Ca^{2+}$, CRT does not support the nuclear export of a hydrophobic NES substrate. The export capacity for a hydrophobic NES substrate can be restored by EGTA-mediated removal of $Ca^{2+}$ from CRT. Under these conditions of low $Ca^{2+}$, RanGTP acts as a stimulatory factor of CRT in a mechanism that involves the assembly of CRT, NES, and RanGTP into a complex (18). Collectively, our data suggest that $Ca^{2+}$ binding determines whether CRT interacts with DBD containing cargo or with NES-containing cargo, and only the latter involves Ran.

The characterization of CRT over the past several years has revealed that it can associate with structurally distinct substrates (20). Studies of the chaperone-like functions of CRT

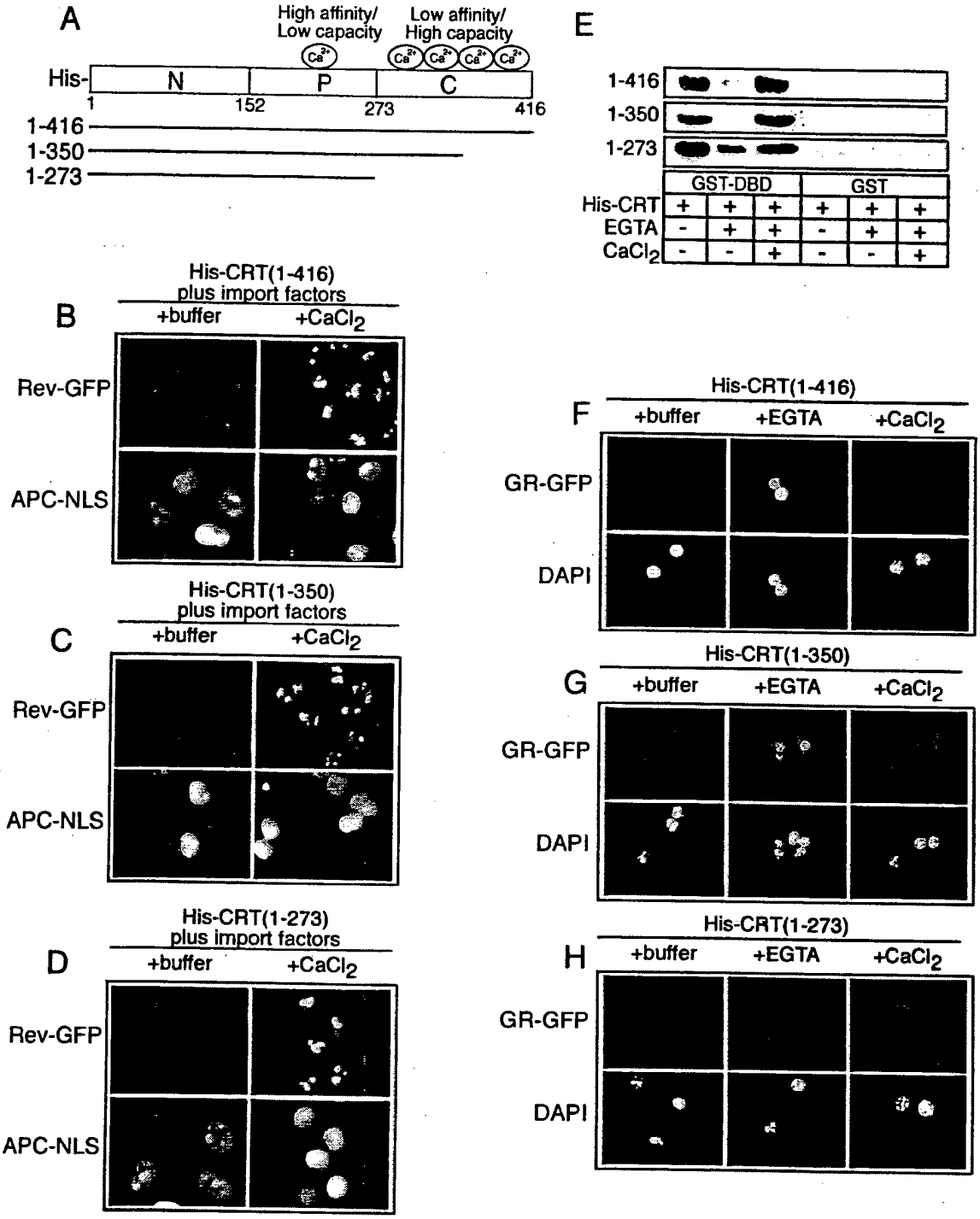



FIG. 6. The low-affinity $Ca^{2+}$ binding sites in the C-terminal domain of CRT are not essential for nuclear export activity. (A) Diagram of the CRT structure and sites of $Ca^{2+}$ binding. (B to D) Export assays performed with CRT proteins containing deletions in the low-affinity $Ca^{2+}$ binding C-terminal domain. CRT lacking the entire C-terminal domain retains its $Ca^{2+}$-dependent inhibition of NES export, indicating that this is probably due to the high-affinity, low-capacity $Ca^{2+}$ binding site. (E) EGTA-sensitive binding of CRT to the DBD is partially lost on removal of the C-terminal domain (residues 1 to 273). (F to H) Removal of the entire C-terminal domain (residues 1 to 273) from CRT abrogates the EGTA-sensitive export of GR. Although the C-terminal domain of CRT is not required for export, it is necessary for $Ca^{2+}$ regulation of GR export.

proposed to occur in the ER lumen have addressed the in vitro interaction of CRT with oligosaccharides, both free in solution and as a structural component of glycoproteins. CRT is capable of binding directly to high-mannose-containing oligosaccharides, and CRT can partially suppress the aggregation of certain glycosylated proteins exposed to elevated temperatures (35). It can also suppress the thermal aggregation of nonglycosylated proteins (35). These data have been taken as evi-

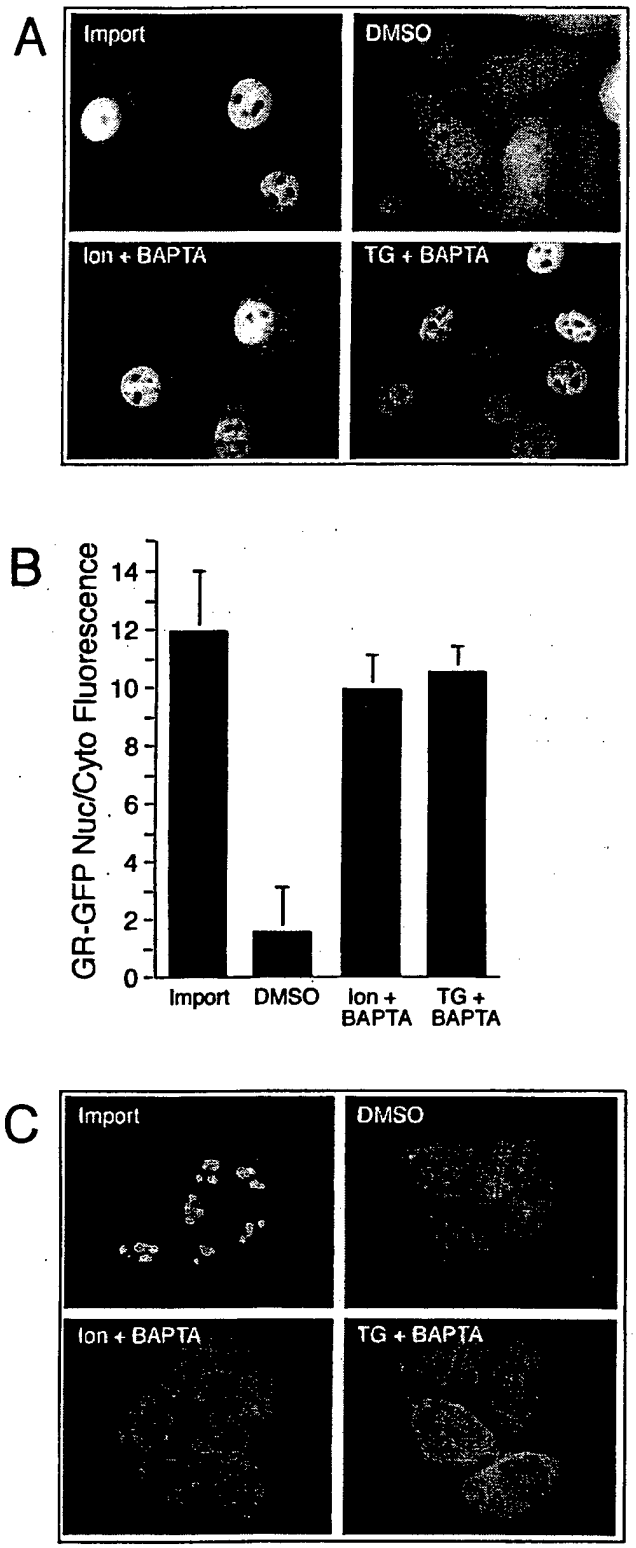



FIG. 7. $Ca^{2+}$ depletion in vivo inhibits the nuclear export of GR. (A) Representative fields of cells expressing GR-GFP after induction of nuclear import with Dex (Import). Following Dex removal, the cells were treated for 5 h with vehicle (DMSO), Ion and BAPTA-AM, or TG and BAPTA-AM, and the GR-GFP distribution was recorded in living cells. (B) Measurements of the nuclear/cytoplasmic (Nuc/Cyto) ratios of GR-GFP fluorescence in cells incubated under conditions that deplete $Ca^{2+}$. Depletion of luminal stores of $Ca^{2+}$ with Ion and TG and chelation with BAPTA-AM results in significant reduction of GR export to the cytoplasm. (C) Depletion of luminal $Ca^{2+}$ stores does not inhibit NES export. The RGG 2.2 cell line expressing Rev-GFP was treated with Dex to induce importing of the reporter, which concentrates in the nucleoli. The cells were then maintained for 6 h under conditions that deplete $Ca^{2+}$. During the last 1 h, Dex was removed to allow Rev-GFP export to the cytoplasm. Rev-GFP export was observed whether or not $Ca^{2+}$ was depleted.

dence that in the ER, CRT may use both its lectin and peptide binding properties to stabilize protein-folding intermediates. The view that CRT participates in the folding of glycoproteins in the ER is logical, given the sequence relatedness (34% identity; expectation value, $<10^{-50}$) and functional similarities to the ER chaperone calnexin. Calnexin is a key component of a quality control pathway that monitors the folding state of glycoproteins in the ER. The current view is that calnexin and CRT interact transiently with protein-folding intermediates, effectively retaining them in the ER until a native conformation is attained (17).

Studies of the functions of CRT that occur outside the ER have yielded an even more diverse collection of substrates (20). These include the cytoplasmic domain of α-integrins (33), steroid hormone receptors (6), NES-containing proteins (18), and a stem-loop structure from rubella virus RNA (38). Some of these substrates contain a binding site for CRT that can be recognized at the sequence level, while other substrates seem structurally distinct. The site within α-integrin recognized by CRT is the sequence KXGFFKR, which is similar to the sequence KVFFKR within the DNA recognition helix of GR and other nuclear receptors. The NESs in PKI (LALKLAGLDIN) and Rev (LPPLERLTL) are similar to each other, but they show little resemblance to the CRT binding sites in α-integrin and the DNA recognition helix. One feature that is shared by most of these signals is the presence of hydrophobic amino acids that, when mutated, abrogates binding to CRT. In the case of the hydrophobic NES, the critical leucines are probably part of an amphipathic helix and the side chains are predicted to physically contact the receptors Crm1 and CRT. This is not the case for the two phenylalanines within the DNA recognition helix that are necessary for binding to CRT (3). The crystal structures of different nuclear receptors reveal that the phenylalanine side chains pack against the core of the DBD, at least when bound to DNA (32). Thus, the FF-to-AA mutations in the DNA recognition helix of GR that inhibit nuclear export probably affect the folding of the DBD, resulting in a structure that is no longer recognized by CRT (3).

How does CRT recognize its substrates? Our present study and previous work from other laboratories suggest that the N-terminal domain mediates substrate recognition and that regulation of binding is imparted by the $Ca^{2+}$-dependent conformational changes in the P-domain and C-terminal domain of CRT. Multiple low-affinity $Ca^{2+}$ binding sites are contained within the C-terminal domain of CRT, and at least one high-affinity $Ca^{2+}$ binding site is contained in the middle domain (P-domain) (2). $Ca^{2+}$ is a cofactor for CRT binding to oligosaccharides, glycoproteins, and certain nonglycosylated proteins in vitro (43). Moreover, in our assays, $Ca^{2+}$ is necessary for the physical interaction between CRT and the DBD as

measured by direct binding and as assayed by nuclear export of GR in permeabilized cells.

$Ca^{2+}$ binding induces structural changes in CRT that dramatically alter its fragmentation pattern that results when it is treated with several different proteases (7). Of particular relevance to our study is the N-terminal 27 kDa of CRT, which is protected from tryptic digestion in the presence of $Ca^{2+}$ but is highly susceptible in the absence of $Ca^{2+}$ (7). This protease-resistant core of CRT retains at least one high-affinity $Ca^{2+}$ binding site from the P-domain, while all of the low-affinity $Ca^{2+}$ binding sites in the C-terminal domain are removed by digestion. We found that a CRT deletion mutant (residues 1 to 273) that contains the protease-resistant core is still functional for nuclear export of GR in permeabilized cells, indicating that the low-affinity $Ca^{2+}$ binding sites are dispensable for our assays that score a positive interaction between CRT and GR. It is interesting that this CRT deletion mutant (residues 1 to 273) shows some $Ca^{2+}$-independent binding to the DBD, although the high-affinity $Ca^{2+}$ binding site is still present in this mutant. The fact that the C-terminal $Ca^{2+}$ binding sites are necessary for EGTA-induced inhibition of CRT binding to the DBD strongly suggests that the C-terminal domain functions in negative regulation of substrate binding. We have also observed that the N-terminal domain of CRT (residues 1 to 150) expressed as a GST fusion protein is sufficient to mediate nuclear export of GR in permeabilized cells (unpublished observations). This is consistent with previous data showing that the N-terminal domain of CRT can block GR binding to the GRE in vitro (6). All available information indicates, therefore, that the N-terminal domain of CRT is the substrate binding domain that physically contacts the DBD of GR. We infer that the N-terminal domain also binds hydrophobic NES substrates, since the DBD and NES display competitive interactions with CRT in our binding assays.

Emerging structural information provides a context for interpreting how $Ca^{2+}$ might regulate CRT activity. Although the atomic structure of CRT has not yet been solved, the structure of the luminal domain of calnexin, which is 35% identical to CRT, has been solved at 2.9 Å resolution (37). The structure of calnexin reveals that it contains a compact globular domain (residues 61 to 262 and 415 to 458) and a large hairpin that forms a highly extended arm (residues 277 to 410). Tandem repeats in the P-domain comprise the arm, which projects 140 Å away from globular domain. The hairpin structure brings the N- and C-terminal domains together to form the globular domain (37). Thus, the close proximity of these domains may help explain how the occupancy of $Ca^{2+}$ binding sites in the C-terminal domain can influence substrate recognition by the N-terminal domain in both calnexin and CRT. The nuclear magnetic resonance spectroscopy structure of the P-domain of CRT (residues 189 to 288) was recently solved, and, like calnexin, this domain forms a highly extended arm (9). The crystal structure of calnexin and the nuclear magnetic resonance spectroscopy structure of the P-domain of CRT were both solved in the presence of $Ca^{2+}$, and so the basis of $Ca^{2+}$-dependent changes in structure awaits further studies. It has been suggested that the P-domain or arm is a $Ca^{2+}$-sensitive protein interaction site for CRT (9).

Does $Ca^{2+}$ regulate the engagement of CRT with nuclear export substrates? Both cytoplasmic and nuclear $Ca^{2+}$ levels can respond to signaling events, and in some cases the modulation of nuclear $Ca^{2+}$ occurs independently of cytoplasmic $Ca^{2+}$ (36). Depending on the system and method of measurement, $Ca^{2+}$ levels in the cell are thought to increase to approximately micromolar concentrations on stimulation (5). This concentration would be predicted to fill the high-affinity $Ca^{2+}$ binding sites on CRT but not the low-affinity sites. It should be noted, however, that the affinity of CRT for $Ca^{2+}$ was measured in vitro. While it is clear that CRT contains at least two classes of binding sites with different affinities for $Ca^{2+}$, demonstrating the $Ca^{2+}$ occupancy of particular sites on CRT in the cell may require the development of fluorescence-based assays that can register $Ca^{2+}$ binding. This approach could be used to determine if $Ca^{2+}$ is constitutively bound to CRT or if $Ca^{2+}$ binds reversibly, thereby acting as a regulator of CRT function. We speculate that transient increases in $Ca^{2+}$ concentration, theoretically, could promote the assembly of CRT-DBD complexes in the nucleus. The complex could then undergo export to the cytoplasm, where the complex is disassembled, which would be favored by low free $Ca^{2+}$ concentrations.

Although the origin of nuclear $Ca^{2+}$ remains an issue of debate, it is clear that processes including transcription are regulated by nuclear $Ca^{2+}$ (16). Several observations are actually consistent with the notion that $Ca^{2+}$ gradients between the nucleus and cytoplasm might be used to promote the assembly and disassembly of CRT complexes. First, the presence of $Ca^{2+}$ ATPases in the nuclear envelope and the lack of $Ca^{2+}$ buffering and sequestration proteins in the nucleus should prolong the availability of free $Ca^{2+}$ (12). Second, the inner membrane of the nuclear envelope contains both major $Ca^{2+}$ release channels, which may be activated independently of channels in the ER membrane that face the cytoplasm (36). Third, morphological data show that the ER forms invaginations that reach deep into the nucleus (13), which could place $Ca^{2+}$ release channels close to sites of CRT-DBD assembly in the nucleoplasm. While it is clear that $Ca^{2+}$ levels in the nucleus do change in response to signaling events, it remains to be established whether $Ca^{2+}$ regulation of CRT is a dynamic process that involves reversible binding or whether $Ca^{2+}$ is constitutively bound and simply maintains the tertiary structure of CRT.

CRT is thought to function in multiple pathways in the cell (23). In the ER, CRT may function as (i) a chaperone that mediates folding by transiently binding to oligosaccharide side chains and peptide domains, (ii) a $Ca^{2+}$ storage protein involved in $Ca^{2+}$ homeostasis, and (iii) a factor that directly regulates the activity of the $Ca^{2+}$ pumps. At the plasma membrane, CRT may modulate (i) cell migration by interacting with adhesion plaque proteins and (ii) cell-cell interactions by interacting with cell surface proteins. In the nucleocytoplasmic compartment, CRT can regulate (i) the activity of steroid receptors through direct binding to the DBD of these proteins and (ii) the redistribution of steroid receptors to the cytoplasm. Because most of these functions mentioned above have been linked to the $Ca^{2+}$ binding activity of CRT, conditions that influence the $Ca^{2+}$ levels are predicted to regulate the activity of CRT. In this regard, our findings that $Ca^{2+}$ is necessary for the nuclear export function of CRT in vitro and that $Ca^{2+}$ depletion results in GR export defects in vivo suggest that this

pathway is used to regulate the nucleocytoplasmic distribution of GR and possibly other steroid hormone receptors. This link between $Ca^{2+}$ and GR export could negatively regulate transcription through a nuclear transport pathway, perhaps using $Ca^{2+}$ as a second messenger.

## ACKNOWLEDGMENTS

We thank Gordon Hager, Dona Love, and John Hanover for cell lines and Ian Macara, Dirk Görlich, and Larry Gerace for plasmids.

Financial support was provided by the NIH (GM58639-01 to B.M.P.) and the DOD (DAMD 17-00-1-0048 to F.R.).

## REFERENCES

1. **Askjaer, P., T. H. Jensen, J. Nilsson, L. Englmeier, and J. Kjems.** 1998. The specificity of the CRM1-Rev nuclear export signal interaction is mediated by RanGTP. J. Biol. Chem. **273**:33414–33422.
2. **Baksh, S., and M. Michalak.** 1991. Expression of calreticulin in *Escherichia coli* and identification of its $Ca^{2+}$ binding domains. J. Biol. Chem. **266**: 21458–21465.
3. **Black, B. E., J. M. Holaska, F. Rastinejad, and B. M. Paschal.** 2001. DNA binding domains in diverse nuclear receptors function as nuclear export signals. Curr. Biol. **11**:1749–1758.
4. **Black, B. E., L. Lévesque, J. M. Holaska, T. C. Wood, and B. M. Paschal.** 1999. Identification of an NTF2-related factor that binds Ran-GTP and regulates nuclear protein export. Mol. Cell. Biol. **19**:8616–8624.
5. **Brini, M., and E. Carafoli.** 2000. Calcium signalling: a historical account, recent developments and future perspectives. Cell Mol. Life Sci. **57**:354–370.
6. **Burns, K., B. Duggan, E. A. Atkinson, K. S. Famulski, M. Nemer, R. C. Bleackley, and M. Michalak.** 1994. Modulation of gene expression by calreticulin binding to the glucocorticoid receptor. Nature **367**:476–480.
7. **Corbett, E. F., K. M. Michalak, K. Oikawa, S. Johnson, L. D. Campbell, P. Eggleton, C. Kay, and M. Michalak.** 2000. The conformation of calreticulin is influenced by the endoplasmic reticulum lumenal environment. J. Biol. Chem. **275**:27177–27185.
8. **Dedhar, S., P. S. Rennie, M. Shago, C. Y. Hagesteijn, H. Yang, J. Filmus, R. G. Hawley, N. Bruchovsky, H. Cheng, R. J. Matusik, and V. Giguere.** 1994. Inhibition of nuclear hormone receptor activity by calreticulin. Nature **367**: 480–483.
9. **Ellgaard, L., R. Riek, T. Herrmann, P. Peter Güntert, D. Braun, A. Helenius, and K. Wüthrich.** 2001. NMR structure of the calreticulin P-domain. Proc. Natl. Acad. Sci. USA **98**:3133–3138.
10. **Fischer, U., J. Huber, W. C. Boelens, I. W. Mattaj, and R. Luhrmann.** 1995. The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. Cell **82**:475–483.
11. **Fornerod, M., M. Ohno, M. Yoshida, and I. W. Mattaj.** 1997. CRM1 is an export receptor for leucine-rich nuclear export signals. Cell **90**:1051–1060.
12. **Fox, J. L., A. D. Burgstahler, and M. H. Nathanson.** 1997. Mechanism of long-range calcium signalling in the nucleus of isolated rat hepatocytes. Biochem. J. **326**:491–495.
13. **Fricker, M., M. Hollinshead, N. White, and D. Vaux.** 1997. Interphase nuclei of many mammalian cell types contain deep, dynamic, tubular membrane-bound invaginations of the nuclear envelope. J. Cell Biol. **136**:531–544.
14. **Gorlich, D., and U. Kutay.** 1999. Transport between the cell nucleus and the cytoplasm. Annu. Rev. Cell Dev. Biol. **15**:607–660.
15. **Greber, U. F., and L. Gerace.** 1995. Depletion of calcium from the lumen of endoplasmic reticulum reversibly inhibits passive diffusion and signal-mediated transport into the nucleus. J. Cell Biol. **128**:5–14.
16. **Hardingham, G. E., S. Chawla, C. M. Johnson, and H. Bading.** 1997. Distinct functions of nuclear and cytoplasmic calcium in the control of gene expression. Nature**385**:260–265.
17. **Helenius, A., T. S. Trombetta, D. N. Hebert, and J. F. Simons.** 1997. Calnexin, calreticulin, and the folding of glycoproteins. Trends Cell Biol. **7**:193–200.
18. **Holaska, J. M., B. E. Black, D. C. Love, J. A. Hanover, J. Leszyk, and B. M. Paschal.** 2001. Calreticulin is a receptor for nuclear export. J. Cell Biol. **152**:127–140.
19. **Holaska, J. M., and B. M. Paschal.** 1998. A cytosolic activity distinct from Crm1 mediates nuclear export of protein kinase inhibitor in permeabilized cells. Proc. Natl. Acad. Sci. USA **95**:14739–14744.
20. **Johnson, S., M. Michalak, M. Opas, and P. Eggleton.** 2001. The ins and outs of calreticulin: from the ER lumen to the extracellular space. Trends Cell Biol. **11**:122–129.
21. **Katagiri, Y., K. Takeda, Z. X. Yu, V. J. Ferrans, K. Ozato, and G. Guroff.** 2000. Modulation of retinoid signalling through NGF-induced nuclear export of NGFI-B. Nat. Cell Biol. **2**:435–440.
22. **Kehlenbach, R. H., A. Dickmanns, A. Kehlenbach, T. Guan, and L. Gerace.** 1999. A role for RanBP1 in the release of CRM1 from the nuclear pore complex in a terminal step of nuclear export. J. Cell Biol. **145**:645–657.
23. **Krause, K. H., and M. Michalek.** 1997. Calreticulin. Cell **88**:439–443.
24. **Lindsay, M. E., J. M. Holaska, K. Welch, B. M. Paschal, and I. G. Macara.** 2001. Ran-binding protein 3 is a cofactor for Crm1-mediated nuclear protein export. J. Cell Biol. **153**:1391–402.
25. **Liu, J., and D. B. DeFranco.** 2000. Protracted nuclear export of glucocorticoid receptor limits its turnover and does not require the exportin 1/CRM1-directed nuclear export pathway. Mol. Endocrinol. **14**:40–51.
26. **Love, D. C., T. D. Sweitzer, and J. A. Hanover.** 1998. Reconstitution Of HIV-1 Rev nuclear export-independent requirements for nuclear import and export. Proc. Natl. Acad. Sci. USA **95**:10608–10613.
27. **McNally, J. G., W. G. Muller, D. Walker, R. Wolford, and G. L. Hager.** 2000. The glucocorticoid receptor: rapid exchange with regulatory sites in living cells. Science**287**:1262–1265.
28. **Nakielny, S., and G. Dreyfuss.** 1999. Transport of proteins and RNAs in and out of the nucleus. Cell **99**:677–690.
29. **Ostwald, T. J., and D. H. MacLennan.** 1974. Isolation of a high affinity calcium-binding protein from sarcoplasmic reticulum. J. Biol. Chem. **249**: 974–979.
30. **Paschal, B. M., and L. Gerace.** 1995. Identification of NTF2, a cytosolic factor for nuclear import that interacts with nuclear pore complex protein p62. J. Cell Biol.**129**:925–937.
31. **Picard, D., and K. R. Yamamoto.** 1987. Two signals mediate hormone-dependent nuclear localization of the glucocorticoid receptor. EMBO J. **6**:3333–3340.
32. **Rastinejad, F.** 2001. Retinoid X receptor and its partners in the nuclear receptor family. Curr. Opin. Struct. Biol. **11**:33–38.
33. **Rojiani, M. V., B. B. Finlay, V. Gray, and S. Dedhar.** 1991. In vitro interaction of a polypeptide homologous to human Ro/SS-A antigen (calreticulin) with a highly conserved amino acid sequence in the cytoplasmic domain of integrin alpha subunits. Biochemistry **30**:9859–9866.
34. **Rubio, I., P. Buckle, H. Trutnau, and R. Wetzker.** 1997. Real-time assay of the interaction of a GST fusion protein with a protein ligate using resonant mirror technique. Bio/technology **22**:269–271.
35. **Saito, Y., Y. Ihara, M. R. Leach, M. F. Cohen-Doyle, and D. B. Williams.** 1999. Calreticulin functions in vitro as a molecular chaperone for both glycosylated and non-glycosylated proteins. EMBO J. **18**:6718–6729.
36. **Santella, L., and K. Kyozuka.** 1997. Effects of 1-methyladenine on nuclear calcium transients and meiosis resumption in starfish oocytes are mimicked by the nuclear injection of inositol 1,4,5-triphosphate and cADP-ribose. Cell Calcium **22**:11–20.
37. **Schrag, J. D., J. J. Bergeron, Y. Li, S. Borisova, M. Hahn, D. Y. Thomas, and M. Cygler.** 2001. The structure of calnexin, an ER chaperone involved in quality control of protein folding. Mol. Cell **8**:633–644.
38. **Singh, N. K., C. D. Atreya, and H. L. Nakhasi.** 1994. Identification of calreticulin as a rubella virus RNA binding protein. Proc. Natl. Acad. Sci. USA**91**:12770–12774.
39. **Stade, K., C. S. Ford, C. Guthrie, and K. Weis.** 1997. Exportin 1 (Crm1p) is an essential nuclear export factor. Cell **90**:1041–1050.
40. **Steggerda, S. M., and B. M. Paschal.** 2002. Regulation of nuclear import and export by the GTPase Ran. Int. Rev. Cytol. **217**:41–91.
41. **Stoffler, D., K. N. Goldie, B. Feja, and U. Aebi.** 1999. Calcium-mediated structural changes of native nuclear pore complexes monitored by time-lapse atomic force microscopy. J. Mol. Biol. **287**:741–752.
42. **Strubing, C., and D. E. Clapham.** 1999. Active nuclear import and export is independent of lumenal $Ca^{2+}$ stores in intact mammalian cells. J. Gen. Physiol. **113**:239–248.
43. **Vassilakos, A., M. Michalak, M. A. Lehrman, and D. B. Williams.** 1998. Oligosaccharide binding characteristics of the molecular chaperones calnexin and calreticulin. Biochemistry **37**:3480–3490.
44. **Wen, W., J. L. Meinkoth, R. Y. Tsien, and S. S. Taylor.** 1995. Identification of a signal for rapid export of proteins from the nucleus. Cell **82**:463–473.
45. **Wente, S. R.** 2000. Gatekeepers of the nucleus. Science **288**:1374–1377.
46. **Wiechens, N., and F. Fagotto.** 2001. CRM1- and Ran-independent nuclear export of beta-catenin. Curr. Biol. **11**:18–27.

Molecular Cell, Vol. 11, 1093–1100, April, 2003, Copyright ©2003 by Cell Press

APPENDIX 3

# Structural Basis for Bile Acid Binding and Activation of the Nuclear Receptor FXR

Short Article

**Li-Zhi Mi,[1] Srikripa Devarakonda,[1] Joel M. Harp,[2] Qing Han,[1] Roberto Pellicciari,[3] Timothy M. Willson,[4] Sepideh Khorasanizadeh,[2] and Fraydoon Rastinejad[1,2,*]**

[1]Department of Pharmacology
[2]Department of Biochemistry and Molecular Genetics
University of Virginia Health System
Charlottesville, Virginia 22908
[3]Universita Di Perugia
Departimento di Chimica e Tecnologia del Farmaco
06123 Perugia
Italy
[4]Discovery Research
GlaxoSmithKline
Research Triangle Park, North Carolina 27709

## Summary

**The nuclear receptor FXR is the sensor of physiological levels of enterohepatic bile acids, the end products of cholesterol catabolism. Here we report crystal structures of the FXR ligand binding domain in complex with coactivator peptide and two different bile acids. An unusual A/B ring juncture, a feature associated with bile acids and no other steroids, provides ligand discrimination and triggers a π-cation switch that activates FXR. Helix 12, the activation function 2 of the receptor, adopts the agonist conformation and stabilizes coactivator peptide binding. FXR is able to interact simultaneously with two coactivator motifs, providing a mechanism for enhanced binding of coactivators through intermolecular contacts between their LXXLL sequences. These FXR complexes provide direct insights into the design of therapeutic bile acids for treatment of hyperlipidemia and cholestasis.**

## Introduction

Bile acids are the endogenous ligands for FXR, the nuclear receptor that transcriptionally regulates their production, movement, and absorption (Chawla et al., 2001; Goodwin and Kliewer, 2002; Lu et al., 2001). As the end products of hepatic cholesterol catabolism, bile acid production is intimately linked to cholesterol homeostasis. In humans, inherited genetic defects associated with the impairment of bile acid homeostasis lead to severe cholestatic disease and vitamin absorption deficiency (Achord, 1990; Hofmann, 1999a, 1999b). Due to its role in the regulation of cholesterol and bile acid levels, FXR is a potential drug target for treatment of hyperlipidemia and cholestatic disease. To understand the mechanism of bile acid binding and activation in FXR, we solved two crystal structures of its ligand binding domain (LBD) in the presence of a GRIP-1 coactivator peptide. One structure contains the high-affinity bile acid 6α-ethyl-chenodeoxycholic acid (6ECDCA) solved at 2.5 Å resolution, and the second structure contains 3-deoxyCDCA, solved at 2.9 Å resolution (Table 1). Together, these structures show how the physiological activities of bile acids are dictated by the chemical composition of their unique steroid skeleton.

*Correspondence: fr9c@virginia.edu

## Results and Discussion

### Overall Architecture of FXR

The FXR protein was expressed from *E. coli* with a N-terminal polyhistidine tag and purified as described in the Experimental Procedures. Crystals of FXR bound to 6ECDCA or 3-deoxyCDCA were obtained under the same conditions. These crystals exhibited identical crystallographic parameters and showed the same molecular arrangements of FXR (Table 1). The asymmetric units contain two representations of the FXR, herein referred to as the *a* and the *b* complexes, which differed in the number of coactivator peptides bound to FXR (one peptide in the *a* complex, two peptides in the *b* complex). Figures 1–3 show the overall disposition of the LBD, the ligand and coactivator peptides from the 6ECDCA crystal. The electron density maps clearly show that the A ring of the bile acid faces the C-terminal helix 12, the activation function 2 (AF2) of the receptor (Figures 1 and 2A). This orientation was unanticipated, as progesterone, estrogen, testosterone, and glucocorticoids are all oriented in the opposite direction with their D rings facing helix 12 of their respective receptors (Bledsoe et al., 2002; Brzozowski et al., 1997; Sack et al., 2001; Shiau et al., 1998; Williams and Sigler, 1998). Bile acid binding directly places helix 12 against helices 3, 4, and 10 (Figure 1). This receptor conformation represents the "activated" state, whereby helix 12 stabilizes the binding of coactivator peptide (Figures 1A and 3A).

### Bile Acid Binding and Discrimination

Bile acids are water-soluble amphipathic steroids that play important roles in cholesterol disposal and intestinal lipid absorption through micelle formation. The physiochemical properties of the bile acid skeleton are essential in this regard, displaying a convex hydrophobic face and a concave hydrophilic face. The steroid nucleus in bile acids is not flat, due to a hydrogen atom at C5 oriented in the β configuration (Figures 1C and 2A). This substituent causes the A/B ring juncture to be *cis*, forcing ring A to lie outside of the plane of the BCD ring system. As a result, the separation between the 3α-OH and the C-24 carboxylate of CDCA (14 Å) is substantially shorter than the contour length of the molecule (21 Å), giving bile acids a rounded profile that allows a close fit with respect to the pocket in FXR (Figures 1B and 1C). Figures 1B and 1C show how the shape of the FXR ligand binding pocket complements the nonplanar bile acid skeleton.

The FXR ligand binding pocket also utilizes the amphipathic properties of the bile acids to provide additional molecular recognition beyond their unique shape. Bile acids contain distinct α and β faces. The β face is com-

Table 1. Crystallographic Data and Refinement Statistics

| | 6-ethyl-CDCA | 3-deoxy-CDCA |
|---|---|---|
| Space group | $P2_12_12$ | $P2_12_12$ |
| Resolution (Å) | 24.3–2.50 (2.59–2.50) | 15.0–2.90 (3.0–2.9) |
| Unit cell dimension (a, b, c) (Å) | 99.829 107.452 69.280 | 99.535 107.129 69.206 |
| Unique reflections | 24427 | 12668 |
| Completeness (%) | 92.8 (77.2) | 98.9 (99.6) |
| Average redundancy | 4.0 (3.4) | 4.7 (4.8) |
| $R_{sym}$ (%)[a] | 4.4 (28.2) | 7.1 (45.8) |
| $R_{working}$[b] | 25.1 | 25.4 |
| $R_{free}$[c] | 28.1 | 29.0 |
| Rms deviation bonds (Å) | 0.010 | 0.011 |
| Rms deviation angles (deg) | 1.6 | 1.9 |
| Rms dihedral angles (deg) | 19.5 | 22.3 |
| Rms improper angles (deg) | 0.93 | 1.10 |

Values in parentheses are for the highest resolution shells.
[a] $R_{sym} = \Sigma|I_h - <I_h>|/\Sigma I_h$, where $<I_h>$ is the average intensity over symmetry equivalent reflections.
[b] $R_{working} = \Sigma||F_o| - |F_c||/\Sigma|F_o|$, where $F_o$ and $F_c$ are the observed and calculated structure-factor amplitudes.
[c] $R_{free}$ was calculated using 5% of data excluded from refinement.

posed of a continuous hydrocarbon surface punctuated with the angular methyl groups at C18 and C19 (Figure 1C). The primary and secondary bile acids contain a common β face, and a distinct α face on which hydroxyls are displayed in the 3α, 7α, and 12α positions. The relative affinities of chenodeoxycholic acid (CDCA), deoxycholic acid (DCA), cholic acid (CA), and lithocholic acid (LCA) for FXR (Makishima et al., 1999; Parks et al., 1999; Wang et al., 1999) are dictated by their specific pattern of axial hydroxyl groups at the 7 and 12 positions. The side chain oxygen of Tyr366 hydrogen bonds with the axial 7α-OH on ring B, and the lack of this hydroxyl in LCA lowers its affinity. Moreover, as the pocket does not provide any polar side chains to accommodate the 12α-OH group on ring C, the binding of CA and DCA would be energetically costly in comparison to CDCA. The FXR structure also reveals pockets that are not entirely filled by substituents found on naturally occurring bile acids (Figure 1B and 1C). The semisynthetic bile acid 6ECDCA (Pellicciari et al., 2002) places the 6α-ethyl group into one such hydrophobic cavity that exists between the side chains of Ile359, Phe363, and Tyr 366, accounting for its higher affinity.

The current structures also provide an explanation for how conjugated bile acids bind and activate FXR. Mammalian and nonmammalian bile acids are often amidated with glycine or taurine at the C24 carboxylate, and these conjugations are known to only modestly affect the binding affinity and activation efficacy (Makishima et al., 1999; Parks et al., 1999; Wang et al., 1999). The structure shows that the carboxylate oxygens of CDCA hydrogen bond with the guanidino group of Arg328 side chain located at the entry point of the pocket. The proximity of the carboxylate on the ligand to the solvent suggests that conjugated amino acids would be positioned completely out of the pocket and solvent exposed, thus not impacting receptor activation. Since amidation still preserves a carbonyl oxygen for hydrogen bonding with Arg328, the bile acid binding affinity would also remain largely unchanged.

As a sensor of physiological levels of enterohepatic bile acids, FXR binds to its cognate ligands with ~1000-fold weaker affinity than other steroid receptors bind to their cognate hormones, yet FXR shows remarkable specificity for bile acids (Makishima et al., 1999; Parks et al., 1999; Wang et al., 1999). Figures 1B and 1C show that the steroid skeleton and the ligand pocket lack a sufficient number of hydrogen bond donors and acceptors to fully account for bile acid discrimination from other steroid hormones. Instead, the FXR ligand binding pocket has evolved to recognize bile acids both using their unique nonplanar shapes and amphipathic physiochemical properties.

The physical-shape discrimination mechanism employed by FXR for ligand binding precludes essentially all other steroids and cholesterol metabolites from acting on this receptor. Estrogens have an aromatic ring A and thus no possibility of *cis-trans* isomerism at carbon-5. Progesterone, aldosterone, glucocorticoids, and testosterone contain a 4-ene structure in ring A, and cholesterol contains a 5-ene in ring B, and therefore all lack the necessary 5β hydrogen required to juncture A/B rings into *cis*. In addition to their differing shapes, none of these other molecules share the amphipathic and detergent-like properties uniquely associated with the α and β faces of bile acids that are both critical to their biological function and make their binding to FXR possible.

### π-Cation Receptor Activation Trigger

All naturally occurring bile acids contain a *cis*-oriented A/B ring juncture and a 3α-hydroxyl group in their A ring. The FXR structure shows the A ring and the 3-hydroxy group oriented toward helix 12 where they interact with the His444 on helix 10/11 and Trp466, the only residue from helix 12 in close proximity to the ligand (Figure 2A). These two residues function as the receptor's activation trigger through formation of a desolvated π-cation interaction directly contacted and stabilized by the bile acid (Figure 2A). The bile acid A ring also contacts Tyr358, which further restricts the position of His444. The π-cation interaction between the indole ring of a Trp (π electron system) and the Nε (cation) on a perpendicularly oriented His side chain has been described in other biological systems but never observed as a molecular switch for receptor activation (Ma and Dougherty, 1997). We infer that in the absence of ligand, the same His-Trp interaction cannot be stabilized due to insufficient physical

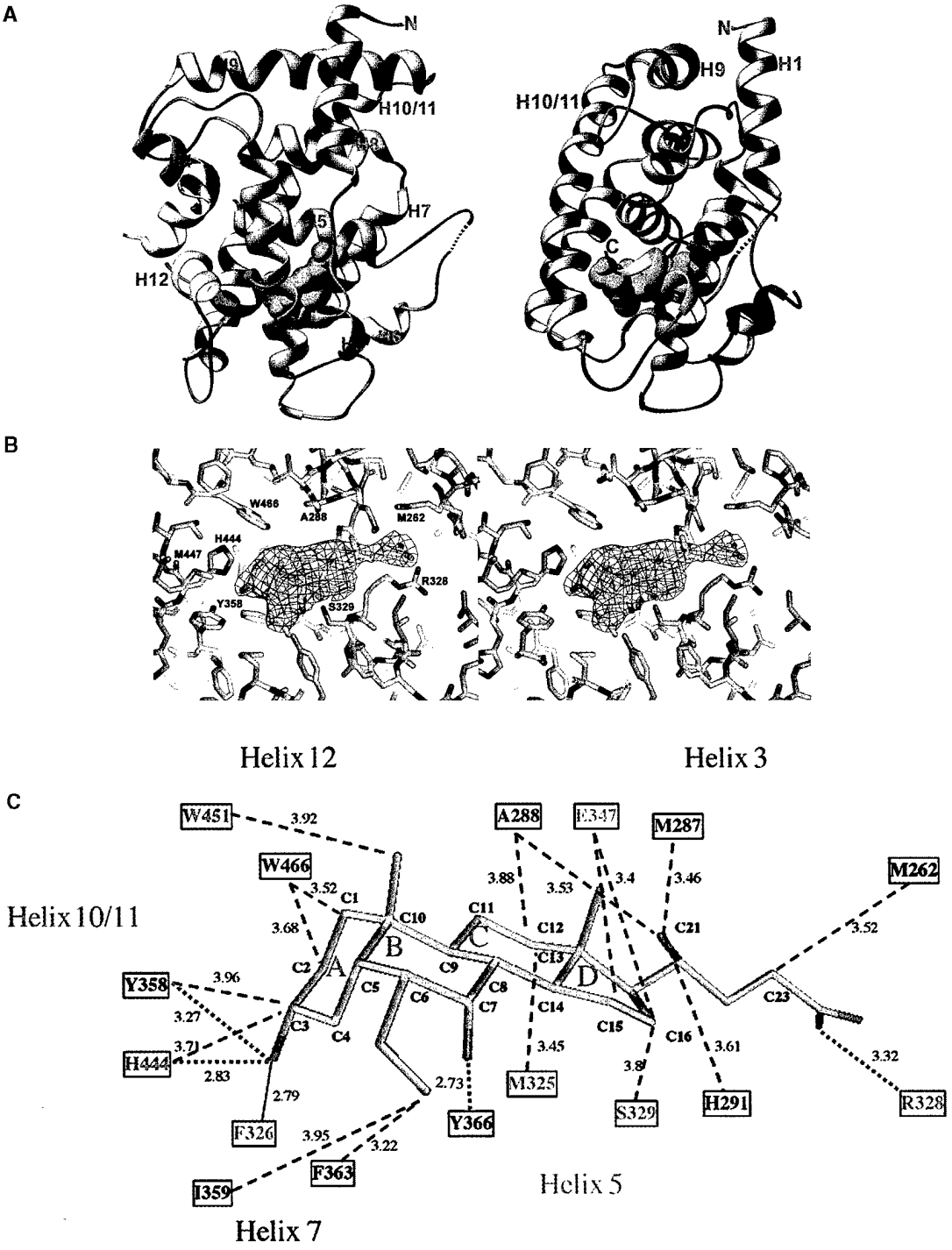



Figure 1. Overall Views of FXR and Its Ligand Interactions

(A) Perpendicular views of the FXR LBD (purple) from complex *b* of the 6ECDCA complex. Helix 12 of the receptor is shown in yellow, the GRIP-1 peptides are shown in blue and red, and 6ECDCA is shown in green.

(B) Stereo view of a simulated annealing omit electron density (Fo-Fc) map showing the bound 6ECDCA. The ligand was excluded in map calculation.

(C) Schematic representation of the FXR/6ECDCA interactions. Dotted red lines indicate hydrogen bonds; dotted blue lines indicate van der Waals contacts.

constraints on these residues. The loss of their interaction would remove the necessary support for helix 12 in its active position.

The 3α hydroxyl on ring A is the most polar functionality on the steroid nucleus due to its equatorial position with respect to the hydrocarbon ring system (Figure 1C). To gauge the relative contributions to receptor activation from the 3-OH group versus the ring A surface, we studied 3-deoxyCDCA in biochemical assays and by X-ray crystallography. We found that the 3-deoxyCDCA

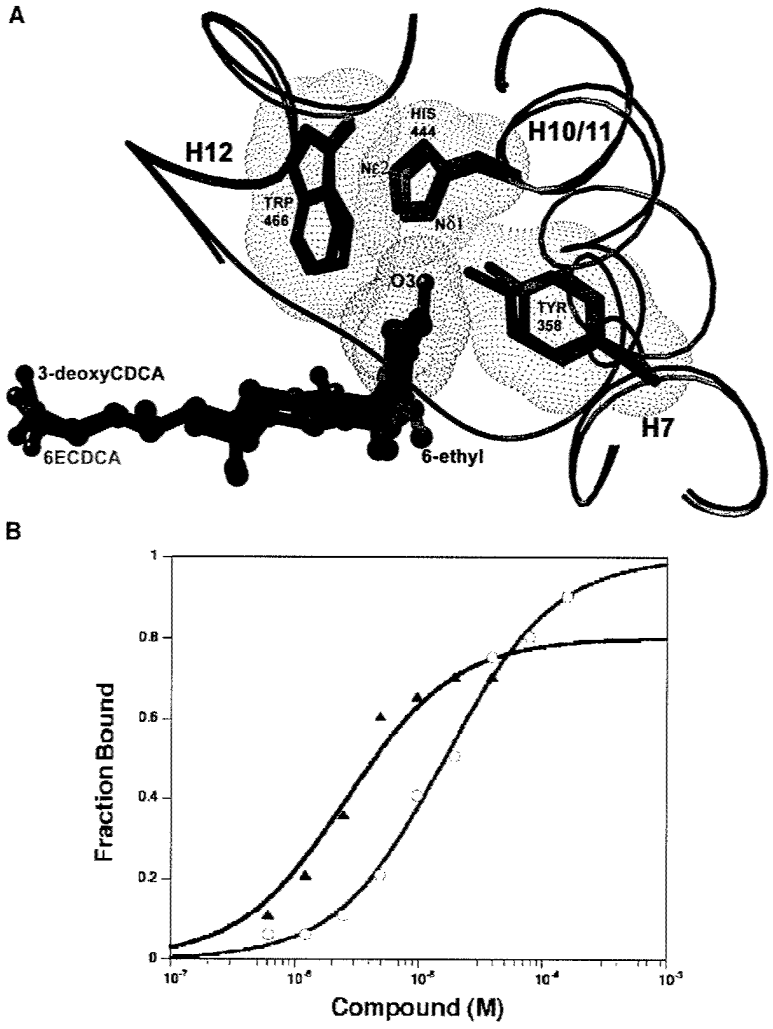



Figure 2. The Activation Mechanism of FXR

(A) 6ECDCA and 3-deoxyCDCA interactions with the activation switch of FXR. Shown is the superposition of their two crystal structures in the vicinity of the ligand and residues His444 and Trp466.

(B) The binding of GRIP-1 NID-3 coactivator peptide to FXR as a function of increasing 3-deoxyCDCA (black triangles) and CDCA (red circles).

is able to recruit the region-3 LXXLL motif of GRIP-1 with an $EC_{50}$ of 3 μM, in comparison to CDCA which shows an $EC_{50}$ of 16 μM (see Figure 2B). The small increase in potency is likely due to the decreased hydrophilicity of 3-deoxyCDCA, which aids desolvation of the ligand. These results suggested that the 3-hydroxyl alone is not responsible for activation of the receptor. To directly visualize how the 3-deoxy bile acid activates FXR, we solved the crystal structure of its complex and examined the ligand with respect to the activation trigger (Figure 2A). We found that ring A of the 3-deoxyCDCA still contacts and orients His444 and Trp466, allowing the same π-cation interaction as in the 6ECDCA complex. A hydrogen bond to the side chain nitrogen on His444, no longer available from the 3-OH of ring A, is instead donated by the side chain OH group of Tyr358. Overall, the structure shows helix 12 in the identical activated conformation seen in the 6ECDCA structure and still supporting the association of coactivator peptide. Therefore, a correctly positioned ring A is the dominant factor mediating the agonist function of bile acids.

A sequence alignment shows that the amino acids responsible for the binding of the sex steroids, glucocorticoids, and mineralocorticoids are more closely conserved between their receptors than with FXR (Figure 4). Nevertheless, all of these receptors form their ligand binding pockets largely via determinants located on helices 3, 5, 7, 10/11, and 12. Importantly, every residue responsible for the binding of rat FXR to bile acids is identical in the human FXR protein (Figure 4). However, mouse FXRβ, which is a pseudogene in humans, contains several differences at these residue positions but also appears to use lanosterol as its endogenous ligand (Otte et al., 2003). As the π-cation activation trigger had not been previously seen in other steroid receptors, we compared sequences in Figure 4 to look for the conservation of this pair of His and Trp residues. Besides the FXR proteins, LXRα and LXRβ receptors are the only other receptors with identically positioned His and Trp residues (Figure 4). As LXRα and LXRβ bind to oxysterols and regulate genes that are also related to cholesterol and bile acid homeostasis (Lu et al., 2001; Peet et al., 1998), these receptors may be activated via a similar π-cation molecular switch.

### Coactivator Interactions

The GRIP-1 coactivator contains three separate nuclear receptor interaction domains (NIDs), each consisting of

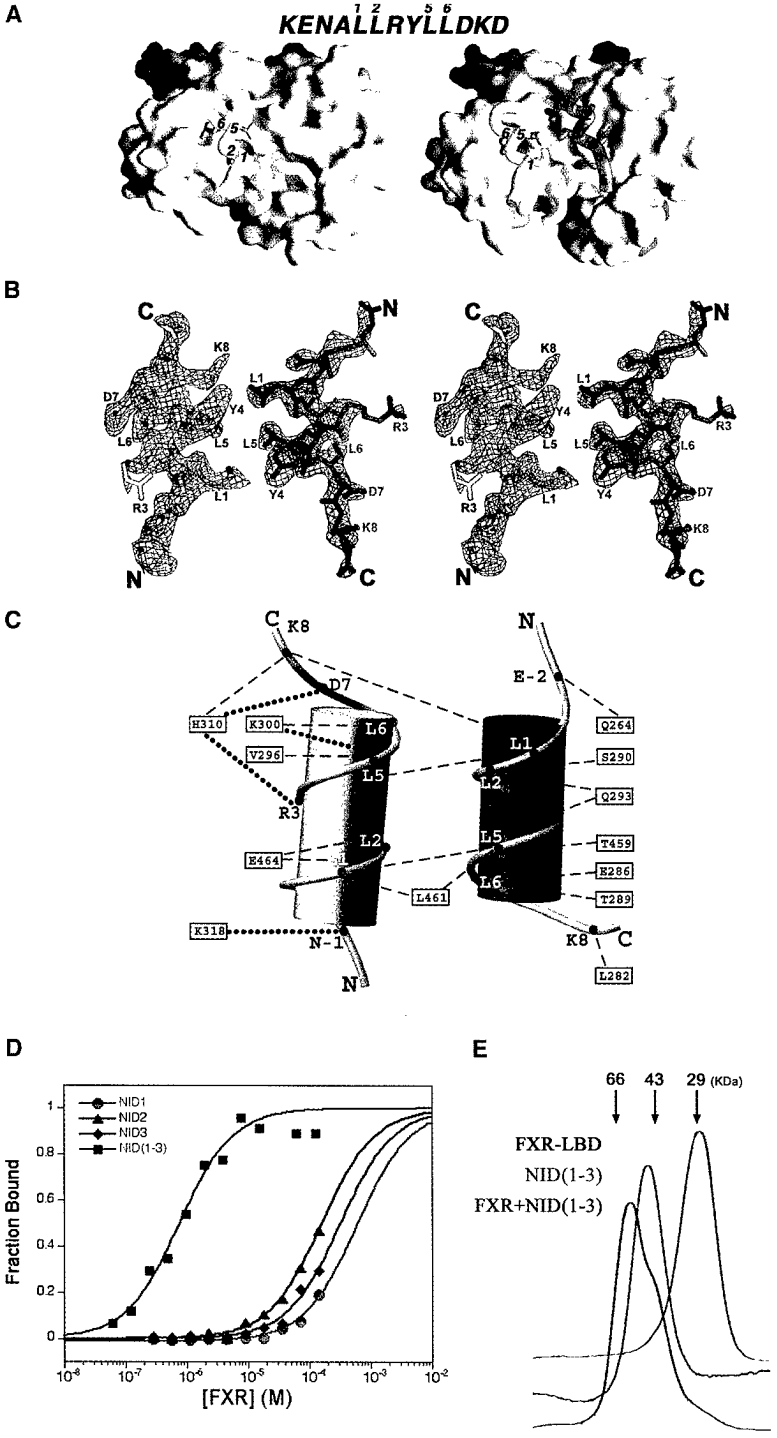



Figure 3. The Interaction of FXR with Coactivator

(A) Comparison of the *a* and *b* complexes in the asymmetric unit of the 6ECDCA structure. Helix 12 is shown in yellow, the peptide in the primary coactivator groove is shown in blue, and the second peptide seen in the *b* complex (of both the 6ECDCA and 3-deoxyCDCA containing structures) is shown in red. The numbers indicate the location of the conserved leucines.

(B) A section of a composite omit map (2Fo-Fc) showing electron density for the two coactivator peptides in complex *b*. The primary coactivator peptide is in blue; the second peptide is in red.

(C) A schematic representation showing the contacts of the primary coactivator peptide (blue) with the second peptide (red) and with FXR (residues indicated in boxes). Circles represent hydrogen bonds; dotted lines indicate van der Waals contacts.

(D) FXR binding to different NID regions of GRIP1 in the presence of CDCA. The NID-1 peptide has the sequence KGQTKLLQLLTTK, the NID-2 is EKHKILHRLLQDS, and the NID-3 is KENALLRYLLDKD. The NID (1-3) protein consists of the three consecutive NID regions of GRIP-1.

(E) Gel filtration chromatography (Pharmacia 100 cm Superdex-75) showing the migration of FXR-LBD, the NID (1-3) region of GRIP-1, and the FXR-LBD + NID (1-3). The elution profiles are shown as a function of increasing time (left to right). The migration of several molecular weight standards is indicated at the top. FXR-LBD elutes at the expected size of a monomer (~29 kDa). The NID (1-3) polypeptide alone elutes at a larger size (~46 kDa) than expected (calculated mass ~24 kDa), presumably due to a noncompact fold. FXR-LBD + NID (1-3) migrates at a size of ~60 kDa, consistent with one FXR-LBD bound to a more compact NID (1-3) and not consistent with the cooperative binding of two FXR-LBDs subunits with one NID (1-3) fragment (expected migration at a size >82 kDa).

a LXXLL motif (Darimont et al., 1998; Xu et al., 1999). Synthetic peptides corresponding to NID-1, NID-2, and NID-3 regions bind to bile acid-activated FXR weakly, with NID-1 binding with $K_D > 500$ μM, NID-2 binding with $K_D$ of ~160 μM, and NID-3 binding with $K_D$ of ~330 μM (Figure 3B). The coactivator peptide in the structures is NID-3, with the sequence KENALLRYLLDKD. Both *a* and *b* complexes show a peptide binding in the primary activator groove as described for other receptors (Figure 3A) (Bledsoe et al., 2002; Brzozowski et al., 1997; Darimont et al., 1998; Gampe et al., 2000; Nolte et al., 1998; Shiau et al., 1998). However, the *b* complexes of both the 6ECDCA and 3-deoxyCDCA structures show a second copy of NID-3 bound adjacent to the primary coactivator peptide (Figure 3A). This arrangement allows the two antiparallel NIDs to form cross interactions via their hydrophobic side chains in a manner that requires the agonist conformation of helix 12. The inability of the second peptide to occupy its site in complex *a* appears to be caused by the lack of the necessary access arising

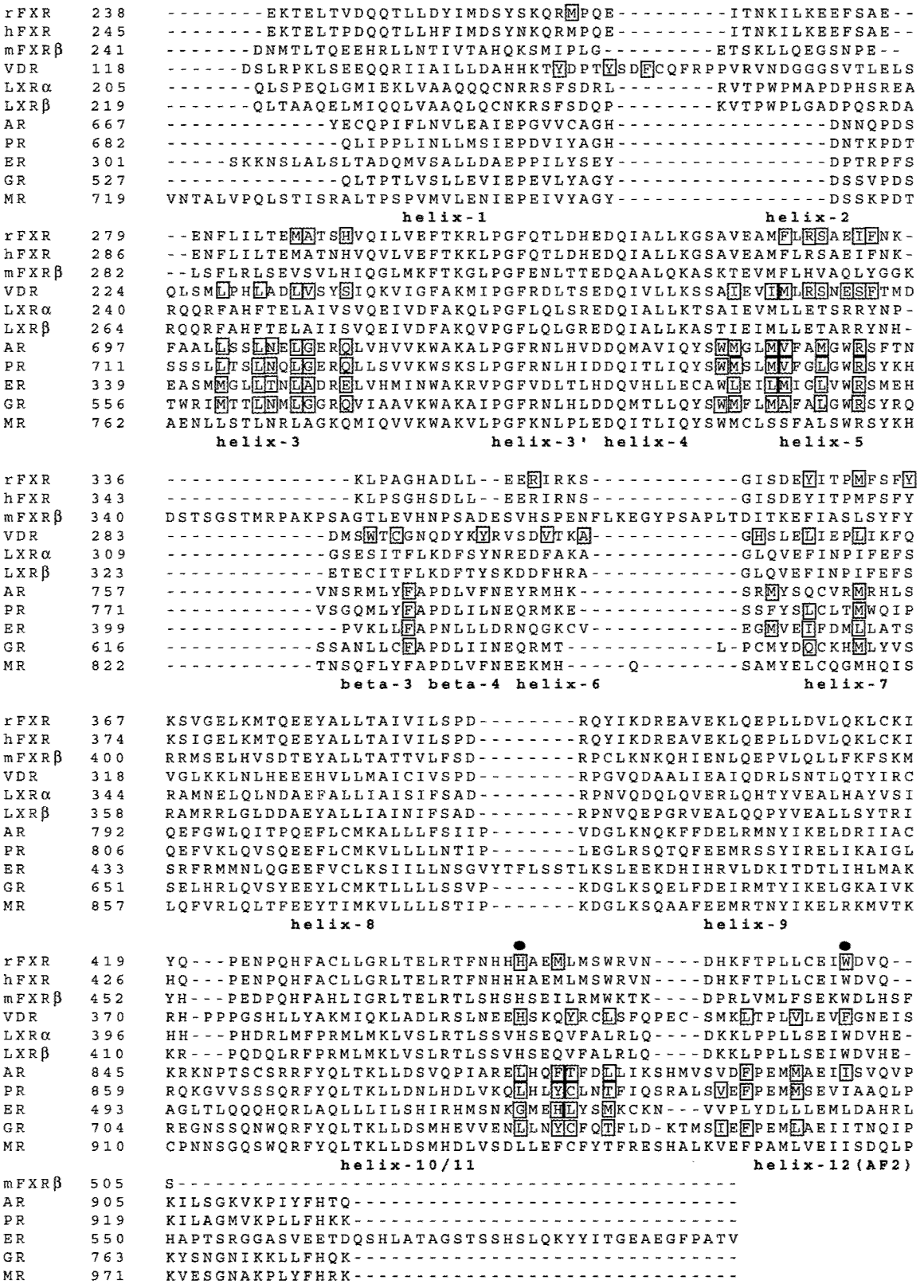



Figure 4. Sequences Used for Ligand Binding by the Steroid Receptors

Yellow shading shows the position of the α helices; blue shows β strands. Red boxes indicate the residues that contact the steroids as determined from crystal structures (Bledsoe et al., 2002; Brzozowski et al., 1997; Rochel et al., 2000; Sack et al., 2001; Shiau et al., 1998; Williams and Sigler, 1998), and black ovals indicate the residues involved in forming the π-cation molecular switch for ligand-induced activation of FXR. All of the residues in rFXR that contact CDCA-derived ligands are identical with human FXR but not as closely shared with mFXRβ, which has been shown to respond to lanosterol (Otte et al., 2003). The crystal structures of hFXR, mFXRβ, LXRα, LXRβ, and MR are unavailable.

from a crystal packing effect involving a symmetry related subunit. Figure 3B shows the electron density maps for both peptides in complex *b*, and Figure 3C shows a schematic drawing of these peptide-peptide and peptide-FXR contacts.

Within the NID-3 LXXLL motif, Arg3 and Tyr4 (occupying the XX positions) are not conserved among the coactivator NIDs and are solvent exposed in the complexes described here (Figure 3A). By contrast, Leu2 and Leu6, which are highly conserved in coactivator NIDS, are responsible for forming the peptides' contacts to FXR (Figure 3C). Leu5 forms the peptide-peptide interactions with Leu1 in the *b* complexes (Figures 3A and 3C). Through the two sets of Leu5/Leu1 interactions, the binding of one peptide enhances the binding of the second peptide by providing a more favorable interaction site. These peptide-peptide interactions would suggest that two NIDs contained within the same polypeptide should bind to FXR with significantly higher affinity, due to the constraints on their physical freedom.

Figure 3B shows that an extended GRIP-1 region (NID [1-3]) that contains more than one NID motif does in fact bind to the FXR more with 100 times greater strength ($K_D$ 0.8 $\mu$M) than any of its isolated NID elements. This higher affinity is consistent with two NIDs in the same coactivator binding cooperatively to a single FXR-LBD subunit. However, another interpretation could be that the larger coactivator fragment is providing the energy for the dimerization of two FXR subunits. The crystal structure of PPAR$\gamma$ in combination with an extended coactivator fragment (Nolte et al.,1998) shows the coactivator forming a bridge across the dimer interface of two LBDs. Therefore, we tested for the subunit dimerization of FXR-LBDs in the absence and presence of the NID (1-3) fragment (Figure 3E). Using gel filtration chromatography, we find that the FXR-LBD is monomeric in solution and that the addition of GRIP-1 NID (1-3) results in a complex consistent in size with a single subunit of FXR-LBD bound to a single NID (1-3) polypeptide. These results, together with the binding data in Figure 3D and the direct crystallographic observations shown in Figures 3A and 3B, are all consistent with a p160 coactivator forming its high-affinity interactions with an FXR-LBD by providing two LXXLL surfaces for the interaction.

## Summary and Conclusions

The current structural study demonstrates a series of unanticipated findings associated with bile acid binding and activation of FXR. The bile acids are discerned from other steroids by their unusual physical shape and amphipathic properties. Bile acids occupy their pocket with their steroid nucleus positioned in the reverse orientation of all other steroids. This arrangement allows the distinctive *cis*-oriented ring A to impart the activated state to FXR by directly contacting a $\pi$-cation molecular switch deep in the pocket. The 3-OH group, which is conserved in all common bile acids, is not necessary for FXR activation. FXR is able to interact with a LXXLL motif that occupies the primary coactivator site, as well as a second motif that requires the agonist conformation of the receptor and further enhances the binding affinity of coactivator. Importantly, none of these features of FXR could have been predicted from structural studies thus far reported for other nuclear receptors.

Therapeutic bile acid supplementation has been a mainstay of Eastern medicine for centuries (Achord, 1990). More recently the bile acid UDCA has been used to treat cholestasis by displacement of more toxic bile acids from the enterohepatic pool (Hofmann, 1999b). However, UDCA is FXR silent and as such does not block de novo bile acid synthesis and has limited clinical efficacy. With the availability of these FXR structures, the potential exists to design new therapeutic bile acids with a modified balance of physiochemical and biological properties for treatment of cholestasis. In addition, the insights onto the molecular details of receptor activation should guide the development of nonsteroidal FXR modulators to impact hepatic cholesterol homeostasis, dietary cholesterol absorption, and serum triglycerides levels.

## Experimental Procedures

### Crystallization and Structure Determination

The ligand binding domain of rat FXR was subcloned into Nde1-BamH1 sites of pET16b vector (Novagen) and expressed with an N-terminal his-tag in BL21-DE3 *E. coli*. The cells were grown overnight in the presence of 0.5 mM IPTG. These cells were pelleted and frozen for long-term storage. For protein purification, the cells were thawed and sonicated in lysis buffer (400 mM NaCl, 10 mM imidazole, 10% glycerol, 20 mM Tris [pH 8.0]). The protein was purified using conventional Ni-NTA chromatography followed by loading on a 100 cm Superdex G200 column (Pharmacia). Ligand was added, and the protein concentration brought to 4–5 mg/ml by ultrafiltration followed by the addition of coactivator peptide. Crystals were grown in hanging drops by using polyethylene glycol as the precipitant, together with ethylene glycol (for cryoprotection) and 0.1M PIPES (pH 6.4). Crystals were flash frozen in liquid nitrogen, and the diffraction data were collected at beamline APS BM-19 in Argonne National Laboratory. Diffraction data were scaled and reduced with HKL2000 (Otwinowski and Minor, 1997). The structure of the 6ECDCA complex was solved using the molecular replacement with MOLREP (Vagin and Teplyakov, 1997). The search model consisted of the RAR LBD (PDB 1lbd) that was altered to remove most of the unconserved loops. The 3-deoxyCDCA structure was subsequently solved with molecular replacement using the refined 6ECDCA coordinates for the search model. Difference (Fo-Fc) maps calculated using phases derived from the 6ECDCA structure, but structure factors from the 3deoxyCDCA data clearly showed strong negative peaks corresponding to the 3-OH and 6-ethyl positions of the 3-deoxyCDCA, consistent with the differences in these ligands. Both structures were refined with multiple rounds of refinement with CNS (Brunger et al., 1998) followed by manual rebuilding.

### Coactivator Binding Assays

The binding of coactivator peptides to FXR was measured using a fluorescence anisotropy-based assay with a Beacon 2000 (PanVera) system. Synthetic peptides were made at the University of Virginia Biomolecular Research Facility, corresponding to sequences shown in the Figure 3B legend. These peptides, as well as *E. coli*-expressed GRIP-1 NID (1-3) protein (sequence 563–767 of GRIP-1/SCR-2), were N-terminal fluorescein-labeled using previously described methods (Jacobs et al., 2001). To measure the $K_D$ values for FXR-coactivator binding, the fluorescence anisotropy values of the peptides (50 nM concentrations) were measured in the presence of 2.5 mM bile acid concentrations and increasing FXR protein concentration. To determine $EC_{50}$ values, the fluorescence anisotropy values of peptide (50 nM) were measured in the presence of FXR (0.1 mM) and increasing ligand concentrations. Data were fitted using the equation $A = [A_f - (A_b - A_f)][\text{protein}]/(K_D + [\text{protein}])$, where $A_f$ and $A_b$ are the anisotropy values of the free and bound peptides, respectively.

### Acknowledgments

We thank Dr. C. Wright, Dr. W. Minor, and S. Jacobs (University of Virginia) for their helpful suggestions and assistance with the

crystallography, and Dr. Alan Hofmann (UCSD) for providing us with 3-deoxyCDCA. J.M.H. is a fellow of the Rett Syndrome Research Foundation. This work was supported by grants from the American Heart Association to S.K. and F.R.

Received: December 13, 2002
Revised: February 4, 2003
Accepted: February 14, 2003
Published: April 24, 2003

## References

Achord, J.L. (1990). Is oriental folk use of bear bile vindicated (yet)? Gastroenterology *98*, 1090–1091.

Bledsoe, R.K., Montana, V.G., Stanley, T.B., Delves, C.J., Apolito, C.J., McKee, D.D., Consler, T.G., Parks, D.J., Stewart, E.L., Willson, T.M., et al. (2002). Crystal structure of the glucocorticoid receptor ligand binding domain reveals a novel mode of receptor dimerization and coactivator recognition. Cell *110*, 93–105.

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. *54*, 905–921.

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

Chawla, A., Repa, J.J., Evans, R.M., and Mangelsdorf, D.J. (2001). Nuclear receptors and lipid physiology: opening the X-files. Science *294*, 1866–1870.

Darimont, B.D., Wagner, R.L., Apriletti, J.W., Stallcup, M.R., Kushner, P.J., Baxter, J.D., Fletterick, R.J., and Yamamoto, K.R. (1998). Structure and specificity of nuclear receptor-coactivator interactions. Genes Dev. *12*, 3343–3356.

Gampe, R.T., Jr., Montana, V.G., Lambert, M.H., Miller, A.B., Bledsoe, R.K., Milburn, M.V., Kliewer, S.A., Willson, T.M., and Xu, H.E. (2000). Asymmetry in the PPARγ/RXRα crystal structure reveals the molecular basis of heterodimerization among nuclear receptors. Mol. Cell *5*, 545–555.

Goodwin, B., and Kliewer, S.A. (2002). Nuclear receptors. I. Nuclear receptors and bile acid homeostasis. Am. J. Physiol. Gastrointest. Liver Physiol. *282*, G926–G931.

Hofmann, A.F. (1999a). Bile acids: the good, the bad, and the ugly. News Physiol. Sci. *14*, 24–29.

Hofmann, A.F. (1999b). The continuing importance of bile acids in liver and intestinal disease. Arch. Intern. Med. *159*, 2647–2658.

Jacobs, S.A., Taverna, S.D., Zhang, Y., Briggs, S.D., Li, J., Eissenberg, J.C., Allis, C.D., and Khorasanizadeh, S. (2001). Specificity of the HP1 chromo domain for the methylated N-terminus of histone H3. EMBO J. *20*, 5232–5241.

Lu, T.T., Repa, J.J., and Mangelsdorf, D.J. (2001). Orphan nuclear receptors as eLiXiRs and FiXeRs of sterol metabolism. J. Biol. Chem. *276*, 37735–37738.

Ma, J.C., and Dougherty, D.A. (1997). The cation-π interaction. Chem. Rev. *97*, 1303–1324.

Makishima, M., Okamoto, A.Y., Repa, J.J., Tu, H., Learned, R.M., Luk, A., Hull, M.V., Lustig, K.D., Mangelsdorf, D.J., and Shan, B. (1999). Identification of a nuclear receptor for bile acids. Science *284*, 1362–1365.

Nolte, R.T., Wisely, G.B., Westin, S., Cobb, J.E., Lambert, M.H., Kurokawa, R., Rosenfeld, M.G., Wilson, T.M., Glass, C.K., and Milburn, M.V. (1998). Ligand binding and co-activator assembly of the peroxisome proliferator-activated receptor-γ. Nature *395*, 137–143.

Otte, K., Kranz, H., Kober, I., Thompson, P., Hoefer, M., Haubold, B., Remmel, B., Voss, H., Kaiser, C., Albers, M., et al. (2003). Identification of farnesoid x receptor beta as a novel mammalian nuclear receptor sensing lanosterol. Mol. Cell. Biol. *23*, 864–872.

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

Parks, D.J., Blanchard, S.G., Bledsoe, R.K., Chandra, G., Consler, T.G., Kliewer, S.A., Stimmel, J.B., Willson, T.M., Zavacki, A.M., Moore, D.D., and Lehmann, J.M. (1999). Bile acids: natural ligands for an orphan nuclear receptor. Science *284*, 1365–1368.

Peet, D.J., Janowski, B.A., and Mangelsdorf, D.J. (1998). The LXRs: a new class of oxysterol receptors. Curr. Opin. Genet. Dev. *8*, 571–575.

Pellicciari, R., Fiorucci, S., Camaioni, E., Clerici, C., Costantino, G., Maloney, P.R., Morelli, A., Parks, D.J., and Willson, T.M. (2002). 6alpha-ethyl-chenodeoxycholic acid (6-ECDCA), a potent and selective FXR agonist endowed with anticholestatic activity. J. Med. Chem. *45*, 3569–3572.

Rochel, N., Wurtz, J.M., Mitschler, A., Klaholz, B., and Moras, D. (2000). The crystal structure of the nuclear receptor for vitamin D bound to its natural ligand. Mol. Cell *5*, 173–179.

Sack, J.S., Kish, K.F., Wang, C., Attar, R.M., Kiefer, S.E., An, Y., Wu, G.Y., Scheffler, J.E., Salvati, M.E., Krystek, S.R., Jr., et al. (2001). Crystallographic structures of the ligand-binding domains of the androgen receptor and its T877A mutant complexed with the natural agonist dihydrotestosterone. Proc. Natl. Acad. Sci. USA *98*, 4904–4909.

Shiau, A.K., Barstad, D., Loria, P.M., Cheng, L., Kushner, P.J., Agard, D.A., and Greene, G.L. (1998). The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen. Cell *95*, 927–937.

Vagin, A., and Teplyakov, A. (1997). MOLREP: an automated program for molecular replacement. J. Appl. Crystallogr. *30*, 1022–1025.

Wang, H., Chen, J., Hollister, K., Sowers, L.C., and Forman, B.M. (1999). Endogenous bile acids are ligands for the nuclear receptor FXR/BAR. Mol. Cell *3*, 543–553.

Williams, S.P., and Sigler, P.B. (1998). Atomic structure of progesterone complexed with its receptor. Nature *393*, 392–396.

Xu, L., Glass, C.K., and Rosenfeld, M.G. (1999). Coactivator and corepressor complexes in nuclear receptor function. Curr. Opin. Genet. Dev. *9*, 140–147.

## Accession Numbers

The Protein Data Bank code for the structure reported in this paper is 1OSV.

## APPENDIX 4

*Full title:* Plasticity of the ecdysone receptor DNA binding domain

*Abbreviated title:* Plasticity of EcRDBD

Marek Orlowski[1], Monika Szyszka[1,4], Agnieszka Kowalska[1], Iwona Grad[1], Anna Zoglowek, Grzegorz Rymarczyk, Piotr Dobryszycki, Daniel Krowarsch[2], Marian Kochman[1], Fraydoon Rastitnejad[3] and Andrzej Ozyhar[1]

[1]Institute of Organic Chemistry, Biochemistry and Biotechnology, Division of Biochemistry, Wroclaw University of Technology, Wybrzele Wyspialskiego 27, 50-370 Wroc_aw, Poland

[2]Laboratory of Protein Engineering, Institute of Biochemistry and Molecular Biology, University of Wroclaw, ul. Tamka 2, 50-137 Wroclaw, Poland

[3] Department of Pharmacology, and Biochemistry and Molecular Genetics, University of Virginia Health Sciences System, Charlottesville, VA 22908-0735, USA

[4] *Present address:* Institut de Genetique et de Biologie Moleculaire, BP10142, 67404 Illkirch Cedex, France

*Corresponding author:* Andrzej Ozyhar
Institute of Organic Chemistry, Biochemistry and Biotechnology,
Division of Biochemistry, Wroclaw University of Technology,
Wybrze_e Wyspia_skiego 27, 50-370 Wroc_aw, Poland;
Phone: 004871 320 63 33; Fax: 004871 328 63 67;
e-mail: andrzej.ozyhar@pwr.wroc.pl

*Key words:* 20-hydroxyecdysone, ecdysone receptor, ultraspiracle, nuclear receptor, *Drosophila*, intrinsically unstructured proteins.

ABSTRACT

Ecdysteroids coordinate molting and metamorphosis in insects *via* a heterodimer of two nuclear receptors, the ecdysone receptor (EcR) and the ultraspiracle (Usp) protein. Here we show how the DNA recognition α-helix and the T-box region of the EcR DNA binding domain (EcRDBD) contribute to the specific interaction with the natural response element and to the stabilization of the EcRDBD molecule. The data indicate a remarkable mutational tolerance with respect to the DNA binding function of EcRDBD. This is particularly manifested in the heterocomplexes formed between the EcRDBD mutants and the wild-type Usp DNA binding domain (UspDBD). Circular Dichroism (CD) spectra and protein unfolding experiments indicate that in comparison to the UspDBD, the EcRDBD is a less structured and less stable protein which uses these properties it shape its heterodimeric complexes on DNA not only with previously characterized partners (EcRDBD, UspDBD), but also with the orphan nuclear receptor DHR38. As such, EcR appears to be an intrinsically unstructured protein with a high degree of intramolecular plasticity. Since recently published structural data indicate the ligand binding domain of the EcR is also characterized by the extreme adaptability, we suggest that plasticity of the EcR domains may be a key factor which allows a single EcR molecule to mediate diverse biological effects.

INTRODUCTION

The 20-hydroxyecdysone (20E) is a steroid hormone that regulates molting, metamorphosis, reproduction and many other developmental processes in insects and in other arthropods. 20E functions through a heterodimeric receptor complex comprised of the ecdysone the receptor (EcR) and the *ultraspiracle* (Usp). Both proteins are members of the nuclear receptors superfamily that includes receptors for steroid and thyroid hormones, and retinoids and vitamin D, as well as many other ligand-inducible transcription factors. Nuclear receptors act by interacting with DNA response elements of target genes known as hormone response elements. Although some receptors can bind as a monomers to a single hexameric motif, most receptors bind as a homo- or heterodimers to HREs composed typically of two core hexameric motifs configured as palindromes, inverted palindromes, or direct repeats. While steroid hormone receptors typically bind to palindromes of 5'-AGA(A/T)CA-3' sequence separated by three nucleotides, nonsteroidal receptors bind preferentially as heterodimers to inherently asymmetric HREs organized as direct repeats. In these cases, the retinoid X receptor (RXR) is the promiscuous partner for different receptor. For the thyroid hormone receptor (TR), the retinoic acid receptor (RAR) or the vitamin D receptor (VDR), heterodimerization with RXR strongly increases the efficiency of DNA binding and transcriptional activity. The same observation was made for *Drosophila* EcR isoforms, EcRA, EcRB1 and EcRB2, each being able to form heterodimers with Usp, which exhibits a strong structural and functional similarity to its mammalian homolog RXR.

Despite the ecdysone receptor's similarities to these receptors, it does not share their preference for binding sites composed of directly repeated half-sites. The most known natural 20E-response elements are highly degenerated imperfect palindromes with a single intervening nucleotide. The best characterized element is the pseudo-palindromic element from the *hsp27* gene promoter ($hsp27_{pal}$), which is composed of two heptameric half-sites

sequences separated by one central base-pair. Analysis of the interaction of the Usp and EcR DNA-binding domains (UspDBD and EcRDBD, respectively) indicated that UspDBD acts as a specific anchor that preferentially binds 5'half-site, and thus is a key factor dictating the polarity (i.e. 5'UspDBD-EcRDBD-3') of the heterocomplex on the $hsp27_{pal}$ element. Interestingly, this arrangement was recently observed in the crystal structure of the UspDBD/EcRDBD complex bound to an idealized element organized as inverted repeat of AGGTCA sequence separated by 1 bp (IR-1). The functional significance of the UspDBD – *hsp27pal* interaction has been supported by the mutational experiments conducted in our laboratory. These studies clearly demonstrated that changes in the UspDBD binding affinity, caused by the substitution of the amino acids in the DNA recognition α-helix by the alanine, are directly reflected in the UspDBD/EcRDBD-$hsp27_{pal}$ interaction pattern.

In this report we illustrate how the DNA recognition α-helix of the EcRDBD contributes to the specific interaction with the $hsp27_{pal}$ and to the stabilization of the EcRDBDprotein fold. In contrast to the UspDBD, EcRDBD exhibits unexpected mutational tolerance in its DNA recognition helix with many mutations still allowing efficient complexation on response elements.. . The mutational tolerance is further seen in the T-box region of the EcRDBD, which readily accepts substitution of almost all amino acids by alanine and results in a general reduction of alpha-helical content in the protein, but still allow DNA binding in distinct complexes.. These properties suggest that the EcRDBD behaves in many ways as an intrinsically unstructured protein characterized by high intramolecular flexibility/plasticity. Interestingly, recently published crystallographic data indicate that the ligand binding domain of the EcR is also characterized by a high degreee of ligand tolerance and adaptability, a property that allows targeting by ligands belonging to distinct chemical classes – including both steroid and non-steroidal agonists. Thus the EcR seems to be the first example of the

nuclear receptor having two flexible domains. The possible biological significance of the this structural plasticity is discussed.

## RESULTS

### *Effects of amino acids substitution in the DNA recognition α-helix of the EcRDBD on the interaction with the $hsp27_{pal}$ element*

To identify the critical amino acids involved in *hsp27* EcRE sequence specific binding by EcRDBD we have substituted each of nine amino acids of the putative DNA recognition α-helix of the EcRDBD with alanine (for more details see Fig.1). The wild-type EcRDBD and alanine point mutants were overexpressed in *E. coli* and purified to homogeneity (not shown). The binding affinities of the DBDs were determined by electrophoretic mobility-shift assays (EMSAs) using a double stranded oligonucleotide containing the original sequence from the *hsp27* gene promoter ($hsp27_{pal}$, see Fig. 1B). As $hsp27_{pal}$ was shown to bind specifically and in a synergistic manner both EcRDBD homodimer and EcRDBD/UspDBD heterodimer (albeit with different affinities), we tested the putative influence of alanine substitutions on homo- and heterodimer interaction. The effects of mutations on EcRDBD homodimer binding are illustrated in Fig. 2A and their quantitative analysis is presented in Fig. 2B. The binding of $hsp27_{pal}$ by homodimers was clearly reduced by alanine substitutions at positions G20, K22, R26, R27, S28 and K31. Among these mutants K22A, R26A, R27A and K31A demonstrated the greatest DNA binding defect, while the homodimerization of G20A and S28A mutants was impaired moderately. In contrast, alanine substitution of G23 and T30 increased affinity of the EcRDBD homodimer for the target sequence. Interestingly, the substitution of E19 by alanine did not appreciably alter the affinity of the EcRDBD. This result was unexpected as E19 corresponds to one of the three receptor-specific amino acids

from the α-recognition helix, referred to as the P-box. The P-box residues have been defined as principal factors that underline the discrimination of different response elements.

Since our previous study has demonstrated that the $hsp27_{pal}$ 5' half-site exhibits substantially higher affinity toward the EcRDBD than the $hsp27_{pal}$ 3' half-site, one can assume that the above results can be ascribed to the interaction of the EcRDBD molecules with the high affinity target half-site. To test this supposition EMSAs were carried out with the derivative of the $hsp27_{pal}$ containing only left half-site ($hsp27_{\Delta R}$, see Fig. 1B). As shown in Fig. 2C and 2D the EcRDBD monomer binding pattern mirrored principally that observed for EcRDBD homodimer. Again K22, R26, R27 and K31 residues play a key role in the specific interaction of the EcRDBD with the $hsp27_{pal}$ element. Their mutation drastically impair binding of the EcRDBD monomer to the 5' half-site and hence the homodimer formation (see Fig. 2A and 2B). On the other site, residues E19, G20, G23, S28 and T30 are not obligatory for binding of the $hsp27_{pal}$ 5' half-site, although their mutation clearly changes DNA-binding affinity of the EcRDBD monomer. Interestingly, the careful comparison of the binding patterns presented in Fig. 2B and 2D indicates that at least for two mutants (G20A and S28A) the magnitude of the EcRDBD homodimer affinity changes is higher than that observed for the monomer. This would indicate that G20 and S28 represent residues that are primarily involved in the binding of the 3'half-site of the $hsp27_{pal}$ element, or in the proper folding of the EcRDBD molecule bound to this part of the element in the EcRDBD homodimer.

To test the contribution of the DNA recognition α-helix in the UspDBD/EcRDBD heterocomplex formation we analyzed interaction of the wild-type UspDBD and EcRDBD mutants with the $hsp27_{pal}$ sequence. The results presented in Fig. 2E and 2F reveal that only R27 residue of the EcRDBD is indispensably for the heterodimer formation. This contrasts with the data for the EcRDBD homodimer (and monomer) where mutations of two additional residues (K22 and R26) were detrimental for $hsp27_{pal}$ binding. Thus, it appears that in the

context of the heterodimer the EcRDBD molecule better tolerates these substitutions. Moreover, the overall magnitude of the observed changes for the affinity of the heterodimers formed by the wild-type UspDBD and the EcRDBD E19A, G20A, G23A, S28A and K31A mutants was lower than for EcRDBD homodimer and monomer (compare Fig. 2F with 2B and 2D).

Together these observations may indicate that amino acid residues from the DNA recognition α-helix of the EcRDBD display different $hsp27_{pal}$ binding characteristics in homo- and heterodimers. This may be also reflection of the remarkable adaptability of the EcRDBD molecule when it forms heterodimer with UspDBD (see results below and Discussion).

*Effects of amino acids substitution in the DNA recognition α-helix on the EcRDBD structure*

We used CD spectroscopy to determine whether the above observed differences in the EcRDBD binding activity arose from loss of the particular amino acid residue (i.e. absence of specific contacts made by the residue) or from perturbation of the higher order protein structure induced by the mutation of the residue. Since the mutated amino acids are located in the region of the EcRDBD which has a α-helical character the far-UV CD spectra were recorded and analyzed. Far-UV CD spectroscopy is especially sensitive to the α-helical structure which usually comprises a substantial fraction of the DBD fold of nuclear receptors. The far-UV CD spectrum of the wild-type EcRDBD shows minima at 222 and 206 nm (see Fig. 3). The first minimum (at 222 nm) is characteristic for CD spectrum of α-helical structure in globular proteins. The minimum at 206 nm corresponds to the second band characteristic for helical proteins (at 208 nm). The blue-shift of the 206 nm minimum and its amplitude, which is larger in comparison with that of globular proteins, may arise from the contribution of conformation other than α-helix. As shown in Fig. 3C, 3E, 3G and 3H the CD spectra of K22A, R26A, S28A and T30A mutants are essentially identical to the spectrum obtained for the wild-type EcRDBD, and so results of the DNA-binding analysis for this

proteins can be interpreted (see Discussion) as a result of loss of contacts of particular amino acid residues with DNA. The intensities of the spectra recorded for other mutants are reduced relative to that of the wild-type EcRDBD, indicating either a decrease of the secondary structure content, or change an equilibrium between folded and unfolded molecules. This would indicate that functional effects observed after mutation of the E19, G20, G23, R27 and K31 residues could be interpreted either as a consequence of structural perturbation introduced in the wild-type EcRDBD or as a result of loss of particular amino acid residues (see Discussion). Interestingly, among these residues only substitution of the R27 residue led to a substantial reduction of the EcRDBD DNA-binding affinity in the homo- and heterodimers. Simultaneously, G20A and especially G23A mutants which along with R27A exhibited the greatest changes in the CD spectra, retained almost full ability to form heterodimers on the $hsp27_{pal}$ element. This again points out exceptional adaptability of the EcRDBD molecule.

*Effects of amino acids substitution in the putative T-box of the EcRDBD on the interaction with the $hsp27_{pal}$ element*

The importance of the EcRDBD C-terminal extension (CTE) sequence for the interaction with the $hsp27_{pal}$ has been stressed by mutational experiments, which demonstrated that deletion of the EcRDBD region corresponding to the so-called A-box (see Fig. 1A) abolished EcRDBD ability to bind DNA and thus to form homo- and heterodimers on the $hsp27_{pal}$ element. The true functional role of the EcRDBD CTE has yet however to be fully appreciated since the crystallographic data allowed to visualize only seven A-box residues exclusively in the EcRDBD-RXRDBD and not in the "wild-type" (i.e. EcRDBD-UspDBD) complex. The remarkable conservation of the CTE sequence among all known EcRs (see Fig. 10) suggests that this region may be crucial for the EcR function. The conservation is particularly evident in the N-terminal region of the CTE, corresponding to the

so-called T-box. With the exception of the *Bombyx mori* EcRDBD, amino acids located in these eight positions are absolutely conserved among all known EcRs (see Fig. 10). Furthermore, analysis of the sequence data of other nuclear receptors indicates that the EcRDBD T-box sequence does not exhibit any significant sequence similarity to other members of the superfamily which highlights its potential importance in response element discrimination. To investigate the role of the T-box sequence for the EcRDBD function we obtained in a homogenous form (not shown) mutated EcRDBD derivatives containing a single alanine residue at eight positions in the T-box region. EMSAs experiments clearly show that only two mutants P68A and E74A display the same $hsp27_{pal}$ binding characteristics as observed for the wild-type EcRDBD, i.e. they form in a synergistic way complexes consisting of two EcRDBD molecules (CII) and the affinity of the monomeric and dimeric complexes observed for the mutants are indistinguishable from that observed for the wild-type EcRDBD. R67A and V71A showed reduced affinity with $hsp27_{pal}$ probe – either as monomers or as dimers. The strongest DNA binding defect was observed when second of the T-box prolines (P73) was replaced by alanine. Interestingly, for one of the T-box alanine substitution variants tested (V72A) the affinities of the monomer and homodimer complexes was substantially increased compared with the wild-type EcRDBD.

Given the putative role of the EcRDBD T-box sequence in forming not only EcRDBD homodimers but also EcRDBD/UspDBD heterodimers, we analyzed the effects on forming heterodimers between the wild-type UspDBD and all EcRDBD derivatives containing alanine substitution in the T-box related sequence. Similarly as it was observed above for DNA binding α-helix mutants, changes in the EcRDBD affinity against $hsp27_{pal}$ are generally reflected in the UspDBD/EcRDBD-$hsp27_{pal}$ interaction (Fig. 5). In particular, the DNA binding defect could be observed for heterodimers containing R67A, E69A, C70A, V71A and particularly P73A mutants, while the heterdimer formed with V72A mutant was able to bind

to DNA with the enhanced affinity. However, that the magnitude of the observed changes was considerably lower than for EcRDBD homodimer and that with one exception (P73), the T-box of EcRDBD readily accepts substitutions of all amino acids, when the EcRDBD is complexed with UspDBD. This observation, along with the CD spectroscopy data (see below) indicating that alanine substitution of some T-box residues induces significant structural perturbations, suggest again that EcRDBD molecule exhibits pliability in contact with its partner (UspDBD).

*Effects of amino acids substitution in the putative T-box on the EcRDBD structure*

Similarly as in the case of DNA binding α-helix mutants we used CD spectroscopy to determine whether substitution of the amino acid residues from the T-box region induces any structural changes in the EcRDBD. As shown in Fig. 6 three of the analyzed mutants (R67A, P68A, E74A) exhibited very similar, if not identical, spectra to that of the wild-type EcRDBD. Two of these proteins (P68A, E74A) exhibited the same $hsp27_{pal}$ binding characteristics as observed for the wild-type EcRDBD, while R67A bound $hsp27_{pal}$ probe with reduced affinity either as EcRDBD homodimers or in UspDBD/EcRDBD heterodimers. CD spectra of four other mutants (E69A, C70A, V71A, V72A) have reduced CD intensity compared to wild-type EcRDBD indicating that in the case of these proteins alanine substitutions lowered the content and possibly altered distribution of the secondary structure elements. Only two of these mutants (V71A, V72A) exhibited substantial changes of the affinity in the functional tests, while two others (E69A, C70A) displayed slightly reduced affinity. Finally, CD spectrum of the P73A mutant highlights again potential distinctiveness of the P73 residue within the EcRDBD T-box sequence. A comparison of the P73A and the wild-type EcRDBD spectra shows that substitution of the P73 with alanine enhances intensity of the spectrum which reflects increase of the secondary structure elements content.

*Properties of the Drosophila EcRDBD mutant containing amino acid residues characteristic for the Bombyx T-box*

As demonstrated in Fig. 10 a primary sequence alignment of EcRDBDs indicates that T-box amino acid residues exhibit remarkable conservation among all analyzed EcR sequences. A surprising originality is the sequence of *Bombyx mori* where the non-conservative substitution of P73 with Q, and conservative substitution of V72 with I can be observed. Interestingly, a P residue is present in the *Bombyx* sequence just downstream of the T-box at position 75 and upstream of the putative A-helix (see Fig. 7A). To investigate whether these substitutions conserve the structural/functional integrity of the T-box region the triple mutant of the *Drosophila* EcRDBD, containing at positions 72, 73, and 75 amino acid residues occurring in *Bombyx* EcRDBD, was obtained (Fig. 7A). As shown in Fig. 7B one of the only characteristics of the mutant which differs from the wild-type *Drosophila* EcRDBD is its slightly reduced ability to form EcRDBD homodimers. The substitution of three *Drosophila* residues with the *Bombyx* counterparts has not resulted in any significant changes in the ability of the mutated EcRDBD to act as a partner for UspDBD (Fig. 7C). Simultaneously, CD spectra of the mutant and the wild-type EcRDBD are virtually superimposable (Fig. 7D). This suggests that no overall structural changes were rendered by the triple mutation. The above results suggest that amino acid substitutions characteristic for *Bombyx mori* EcRDBD conserve the structural/functional integrity of the T-box region, at least in terms of the DNA-binding affinity.

*Flexibility of the EcRDBD*

As demonstrated above many of the EcRDBD mutants, including these with the defined structural changes, exhibit a surprising tolerance for alanine substitution because they are able to form efficiently complexes with $hsp27_{pal}$ in the presence of the UspDBD. This suggests the flexibility and adaptability of the EcRDBD that would allow moulding of this protein around

the DNA and the UspDBD. It has been recently suggested that intrinsically unstructured proteins (IUPs) are inherently flexible and that both their local and global structures can be easily shaped by the environment. Such intrinsic plasticity could allow a single protein to recognize a large number of biological targets without sacrificing its specificity. As discussed above the EcRDBD CD spectrum reflects the contribution of structures other than α-helix – most probably random coil structure, which would represent residues that are not specifically assigned to other than the α-helix pure secondary structures. In contrast the UspDBD CD spectrum displays clear α-helical character (Fig. 8A) and is very similar to the CD spectra observed previously for DBDs of the receptors for estrogen and glucocorticoid. Thus in comparison to its partner (i.e. UspDBD) EcRDBD seems to be more unstructured. To determine if this observation is reflected in the stabilities of the EcRDBD and UspDBD molecules we performed chemical denaturation experiments using guanidinium hydrochloride. Since both DBDs does not contain tryptophan, the unfolding reaction was monitored using intrinsic tyrosine fluorescence. The results presented in Fig. 8B show that EcRDBD is extremely unstable protein. Starting from the lowest denaturant concentrations a gradual decrease in the fraction of the unfolded protein could be observed. To check if this denaturation profile is characteristic only for the *Drosophila* EcRDBD molecules, chemical denaturation experiment was carried out for the overexpressed and purified *Bombyx mori* EcRDBD. As shown in Fig. 8B the denaturation profiles of both proteins are virtually superimposable, which suggests that structural liability seems to be a general feature of the EcRDBDs. Notably, the triple mutant (V72I/P73Q/N75P) described in this study exhibits the same denaturation characteristics (see Fig. 8B). It has been recently reported that truncating the highly positively charged C-terminal extensions of DBDs of glucorticoid and estrogen receptors increased stability of both proteins. To test if the stability of the EcRDBD is dependent on the C-terminal region, the denaturation profile was studied for the EcRDBD

mutant ($EcRDBD_{\Delta A}$) with the deleted C-terminal region (A-box) containing many positively charged amino acid residues. Fig. 8B shows that the denaturation curve of this proteins is very similar to the denaturation curve obtained for the wild-type EcRDBD. This, along with data obtained for other DBDs, would imply that the A-box sequence itself does not destabilize the EcRDBD and that the EcRDBD instability is inherent property of the core part of this protein. In opposition to EcRDBD, the UspDBD molecule is more stable. The UspDBD denaturation curve is clearly shifted to a higher concentrations of denaturant (see Fig. 8B) and shows a two state transition, previously reported for DBDs of mammalian steroid receptors.

To answer the question of whether or not guanidinium hydrochloride-denatured Usp and/or EcR DBDs are able to return to active structure, refolding experiments were performed. After dilution of the unfolded DBDs samples incubated wit 4.0 M guanidinium hydrochloride, their $hsp27_{pal}$ binding affinities were analyzed and compared with the activity of the control samples (Fig. 8C and 8D). The control samples contained the same concentration of the respective DBD and the residual concentration of the denaturant as the diluted samples, however DBDs in the control samples were not exposed to the 4.0 M guanidinium hydrochloride. As shown in Fig. 8C and 8D (lanes 3 and 4) UspDBD recovered at least 90% of the control sample activity. Slightly lower recoveries were observed for homodimeric complexes of *Drosophila* and *Bombyx* EcRDBDs – 79% and 75%, respectively (Fig. 8C and 8D, lanes 6, 7 and 9,10, respectively). Interestingly, heterodimeric complexes formed using non-denatured DBD and renatured partner exhibited almost identical activity (ca. VV% of the activity of the control sample containing non-denatured DBDs), irrespective which of the DBD was renatured (Fig. 8C and 8D, lanes 11-14).

Thus, in contrast to UspDBD EcRDBD seems to be characterized by the low content of the defined secondary structure and by the low stability, which together means that in the DNA-unbound state the EcRDBD is likely to be largely unstructured. Nevertheless, even after

denaturation/renaturation cycle the EcRDBD finds its unique native structure/conformation which manifests itself in the ability to form homo – and heterodimers on $hsp27_{pal}$ element. Interestingly, on the same element EcRDBD is also able to form, in a synergistic manner, complexes with the DBD of the *Drosophila* orphan nuclear receptor DHR38 (DHR38DBD) (Fig. 9). So, in contrast to UspDBD, which on the same element does not form complexes with DHR38DBD (not shown), EcRDBDB molecule exhibits remarkable ability to accept different DBDs as a partners on $hsp27_{pal}$.

## DISCUSSION

Although molecular studies of the EcR and the Usp are much less extensive than those of vertebrate heterodimeric nuclear receptors, it is already clear that the EcR exhibits a unique combination of characteristic and thus holds a exceptional position among the family. For example, unlike the vertebrate steroid receptors EcR by itself cannot bind its ligand, 20E; it must heterodimerize with its partner Usp for the stabilization of a ligand-binding conformation. The other very intriguing feature of the EcR/Usp heterodimer is its preference for response elements arranged as highly degenerated imperfect palindromes with a single intervening nucleotide. This attribute clearly distinguishes the EcR/Usp heterodimer from the vertebrate counterparts which prefer inherently asymmetric DNA binding sites composed of directly repeated half-sites for formation of the complexes characterized by the defined anisotropy. An arrangement in which RXR subunit locates on the 5' site of directly repeated elements, is most often observed with RXR heterodimers. Surprisingly, the EcRDBD/UspDBD exhibits a similar anisotropic arrangement not only on highly degenerated natural $hsp27_{pal}$ element but also on idealized IR-1 element containing symmetric AGGTCA half-sites. In both cases the UspDBD binds 5' half-site of the element while the EcRDBD subunit binds to the 3' half-site. The same is also true when EcRDBD forms heterocomplex

on the IR-1 element with the DBD of the RXR (RXRDBD) – the mammalian homologue of the Usp. In contrast to the vertebrate DBDs forming on the direct repeats, the non-symmetric 'head to tail' heterocomplexes with RXRDBD, the EcRDBD/RXRDBD and the EcRDBD/UspDBD complexes are organized on IR-1 element in a 'head to head' manner, with the same fragments of the EcR and the Usp/RXR DBDs brought to the subunit interface. However, since the regions of EcRDBD are distinct in sequence from their counterparts in the Usp/RXR, the dimer interface has clear non-symmetric character. This contrasts with observation made in the case of vertebrate steroid receptor complexes with the inverted repeat elements, in which 'head to head' organized DBDs use identical residues to form perfectly symmetric interactions. Thus the the EcR receptor, which relies on both inverted repeat target and heterodimerization, has characteristics typical of both steroid and non-steroid receptors, and seems to be a evolutionary link between the vertebrate steroid and non-steroid receptors.
In the experiments reported here we have examined the contributions of the individual amino acids located in two particular EcRDBD regions to the specific interaction with the natural response element and to stabilization of the EcRDBD molecule.

*DNA-recognition α-helix*

Recently published crystallographic analysis of the EcRDBD/UspDBD complex have identified the main "hook' residues within the DNA-recognition α-helix of the EcRDBD. It has been shown that these residues, consisting of E29, G20, K22, R26 and R27, form base specific contacts with the 3' half-site of the IR-1 element. Our results indicate that with one exception (E19) all these residues are also critical when 5' half-site of the natural pseudopalindromic element ($hsp27_{pal}$) is used as a target. Two other residues (S28, K31), which were not identified as hook residues in the EcRDBD/UspDBD/IR-1 crystal, seem to be involved in some $hsp27_{pal}$-specific contacts. Interestingly, the same residues were identified to play a key role in EcRDBD homodimer formation. Our previous study has demonstrated that

the 5' half-site of the $hsp27_{pal}$ element exhibits substantially higher affinity against the EcRDBD than the 3' half-site. Thus, it is quite possible that the results obtained for EcRDBD homodimer reflect primarily the influence of the respective mutations on the EcRDBD molecule interacting with the high-affinity 5' half-site. However in the case of two residues (G20, S28) the magnitude of the EcRDBD homodimer affinity changes was clearly higher than it was observed for monomer, which could indicate that G20 and S28 are main contact residues of the EcRDBD located on the *hsp27pal* 3' half-site.

Results presented in this study demonstrate different impact of the amino acid residues of the DNA-recognition α-helix on EcRDBD interaction with DNA. The greatest DNA binding defects were observed for three residues (K22, R26 and R27), which are very often involved in formation of the specific contacts of different DBDs with DNA. However, the pattern of hydrogen bonds between these residues and respective bases varies and is case-specific. Here, simultaneous analysis of the EMSA and CD data demonstrate for the first time that only K22, and R26 are pure contact residues, which are not involved in the stabilization of the DNA-recognition α-helix and possibly of the whole EcRDBD molecule. This is also true for S28, which has not been defined as a hook residue in the EcRDBD/UspDBD/IR-1 structure. All other residues important for effective interaction with the $hsp27_{pal}$, including R27, are to some extent involved in stabilization of the EcRDBD molecule. Thus, the observed changes in the EcRDBD binding activity cannot be solely attributed to the loss of the specific contacts, due to the mutation of the defined amino acid residue. Further experimental support by crystallographic and/or NMR spectroscopy studies is needed to clarify ultimately the involvement of these residues in the specific contacts with the natural, i.e. asymmetric elements. Of particular interest among all analyzed residues was G23. Its substitution by the alanine led to the most pronounced change in the EcRDBD molecule. However, this change is readily accepted by the EcRDBD. This along with other observations indicating that

magnitudes of the affinity changes against DNA were lowered in the EcRDBD/UspDBD heterocomplex suggests that EcRDBD exhibits exceptional plasticity (see below).

*T-box*

Mutational studies and analysis of the crystallographic structures of the nuclear receptor DNA-binding domains bound to asymmetrical response elements have underlined the functional importance of the so-called C-terminal extension (CTE) of the core DBD. The CTE sequence, consisting of the T-box and the adjacent α-helix (A-box), is a unique and characteristic structural element which mediates for the particular receptor its dimerization on asymmetrical elements and significantly contributes to DNA binding specificity. Recently reported crystal structures of the VDR DBD in complex with response elements from three different promoters support this hypothesis and provide further evidence that the T-box together with α-helix are identity elements defining the conformation of the CTE and thus response element discrimination. In this work we have analyzed the roles of the individual amino acids located in the EcRDBD T-box sequence. Accordingly to the data published for vertebrate heterodimeric receptors T-box, together with adjacent A-box, forms a characteristic structural element, which for the particular receptor mediates its dimerization on asymmetrical elements and considerably contributes to DNA binding specificity. Whereas A-box residues are mainly involved in the specific contacts with the response element, T-box performs simultaneously different functions, including specific base-pairing, forming of the dimmer interface and first of all bringing of the A-box helix in the correct position. We examined the effects of alanine substitutions for eight T-box amino acid residues on DNA binding activity and structure stability of the EcRDBD molecule. As shown in Fig. 10, with exception of two positions in *Bombyx mori*, these amino acids are absolutely conserved among all other EcRs. This conservation cannot be easily explained for two residues (P68, E74) which accordingly to our data are dispensable at least for DNA binding and stabilization

of the structure of the EcRDBD molecule. Interestingly, the R67 seems to be a pure contact residue, not involved in the stabilization of the domain. Since, according to the crystallographic data obtained for EcRDBD/UspDBD heterocomplex on IR-1 element this residue has not been reported to be involved in any contacts with DNA, it is reasonably to assume that our observation is specific for natural response element. All other analyzed residues show evidence of structural function as indicated by the CD spectra changes observed. However, in case of V71A, V72A and P73A mutants involves unambiguous change in the DNA binding affinity of the EcRDBD mono- and homodimer. These residues together wit C70 make up the hydrophobic core of the T-box region and in the EcRDBD/UspDBD complex on IR-1 element are localized in this part of the EcRDBD where dramatic change in the polypeptide direction is observed. Two consecutive turns are responsible for it; β-turn which includes C70, V71, V72, and P73, and γ-turn consisting of P73, E74 and N75. The results presented in this work highlight the distinctiveness of the P73 residue – its mutation has detrimental effect on DNA binding affinity of the EcRDBD homodimers and exhibits the strongest effect on the heterodimer affinity. As can be concluded from the structural data obtained for the EcRDBD/UspDBD bound to IR-1 element the P73 residue is not involved in any interactions important for stabilization of the core DBD, thus it is conceivable that its main task is the proper folding of the EcRDBD CTE. Theoretically, EcRDBD deprived of A-box contains all residues needed for the effective formation of the heterocomplex with UspDBD on the response element. This, at least, can be inferred from the above mentioned structural data. However, biochemical data obtained for such EcRDBD derivative have clearly demonstrated that A-box is indispensable for effective heterodimer formation. Furthermore, structural data obtained for only seven out of 25 A-box residues of the EcRDBD complexed with RXRDBD on IR-1 element suggest that A-box is directed to the 5' half-site. Thus, one could speculate that this part of the EcRDBD is involved in some

additional, not seen in the crystal structure, protein – DNA or protein-protein-DNA interactions. It is quite possible that P73, whose amide and carbonyl groups are involved in crucial hydrogen bonds of β- and γ-turn, is a key residue positioning A-box of the EcRDBD. Similar function could be ascribed to the V72 residue. Accordingly to the structure of the EcRDBD/UspDBD heterocomplex on IR-1, the V72 resides lies outside of the core DBD. However, due to the interaction with the phosphate backbone the V72 is locked in the defined position outside. The function of the V71 seems to be more complicated since according to crystallographic data this residue is involved in the hydrophobic interactions with side chains of the core DBD residues, i.e. V29 and Y13. Uniqueness of the V71, V72 and P73 residues is further supported by the results obtained for two residues (E69 and C70) which are located in the short a-helical region starting the T-box sequence. Although their mutations are detrimental for the EcRDBD structure, only a minute effect on the homodimer binding could be detected.

The ultimate elucidation of the role of the EcRDBD CTE needs further experimental support. In particular, this would involve crystallographic analysis of the EcRDBD/UspDBD heterodimers complexed with the natural response elements, which in contrast to the idealized IR-1 element could promote proper, i.e. ordered, folding of the EcRDBD C-terminal sequence. Nevertheless, results presented in our paper clearly indicate that in contrast to the vertebrate receptors the EcRDBD C-terminal sequence is involved in the stabilization of this domain and in the interaction with the specific DNA sequences.

*Plasticity of the EcRDBD*

As shown above the EcRDBD exhibits unexpected mutation tolerance, which is particularly manifested when DNA-binding affinity of the EcRDBD/UspDBD heterodimers is analyzed. At least in the case of the DNA recognition α-helix this property clearly distinguishes the EcRDBD from the UspDBD. It has been shown previously that UspDBD affinity changes

caused by the substitution of the DNA recognition α-helix residues by the alanine are directly reflected in the UspDBD/EcRDBD – $hsp27_{pal}$ interaction pattern. Our unpublished results indicate that this is also true for the UspDBD T-box alanine mutants. This striking difference in the structural adaptability, which has not been observed yet for any vertebrate nuclear receptor DBDs forming heterodimeric complex, can be explained by the different properties of both DBDs. In contrast to UspDBD, the EcRDBD seems to be more unstructured and less stable, which unveils the absence of cooperative unfolding transitions characteristic of a well-defined (i.e. globular) tertiary folding. These properties, along with the previously published gel filtration data indicating that the apparent molecular mass of the EcRDBD (17.4 kDa) is greater than the predicted value (13.5 kDa), suggest that EcRDBD (or at least some fragments of this domain) is akin toIUPs, which are characterized by an almost complete lack of the folded structure and an extended conformation with high intramolecular flexibility and little secondary structure. An prominent feature of IUPs is that they can adopt different structures upon different stimuli or with different partners, which enables their versatile interaction with different targets. This phenomenon termed binding promiscuity or one-to-many signaling, seems to exists also in the case of the EcRDBD. Our paper provides for the first time experimental evidence that the EcRDBD is able to interact on $hsp27_{pal}$ element not only with the EcRDBD and the UspDBD but also with the DHR38DBD molecule. The biological significance of the last observation is not clear, since it has been reported that the full-length EcR and DHR38 are not capable of binding to *hsp27pal* element. Our results however, clearly demonstrate that isolated DBDs of both receptors can interact on this element effectively, which suggests that under certain conditions formation of the EcR-DHR38 complex on DNA might be possible. Another unique functional faculty of the IUPs is that their open structures is largely preserved when they complex with their target, which provides for a disproportionaly large binding surface and multiple contact points. The existence of the large

DNA-binding surface in the EcRDBD and multiple contact points involved in the interaction have been suggested for the first time by mutational experiments. Recently published structural data of the EcRDBD/UpsDBD complex with IR-1 element fully support this hypothesis and show that EcRDBD footprint on DNA extents well beyond its own AGGTCA site to reach over both its 3'-flanking sequences and a large portion of the UspDBD half-site. In total, the EcRDBD footprint extends over 13bp. According to XXXX the exact nucleotide sequence of the response elements affects not only the overall affinity of the receptor for its site but also influences the tertiary structure of the receptor. Therefore site-specific DNA binding may be a trigger for an active intramolecular event that changes the shape of the DBD and in consequence of the whole receptor in a response element-specific way. This could result in binding certain ancillary proteins in a site-specific way. Since transcriptional regulation for a specific gene depends upon interactions of these proteins, the exact DNA sequence of the response elements may be the crucial regulatory factor. The EcRDBD/UspDBD heterodimer complexed with the respective response element seems to be the appropriate candidate for such sequence-dependent ancillary protein selection. Firstly, the integral part of the heterodimer is the EcRDBD, which due to natively unstructured proteins-like properties could easily accommodate DNA-induced changes in the secondary and tertiary structure. Secondly, as demonstrated by mutational studies only XXbp among 15bp in $hsp27_{pal}$ are indispensable for effective interaction either with the complex formed by the full-length EcR and Usp or with their DBDs. Other positions are not obligatory for heterodimerization, and as previously suggested, these positions may be used for the fine tuning of the EcRDBD/UspDBD structure and thus biological response. Since even subtle structural changes in the DBD can have long-range functional consequences it is possible that in the full-length EcR DNA binding could affect tertiary structure of other EcR domains. The EcR ligand binding domain looks like a appropriate candidate for such intramolecular

"crosstalk". According to mutational and crystallographic studies this domain is also characterized by the extreme flexibility and adaptability, properties that allow moulding of the domain around two completely different ligands – steroid and non-steroid agonists. Thus, plasticity seems to be general feature of the EcR molecule. One of the reasons why EcR, the only known ligand-dependent nuclear receptor in insects, has gained such unusual characteristic among all other nuclear receptors may be its key position in the mediation of the 20E signal transduction pathways. The insect haemolymph carries a wide range of endogenous ecdysteroids, some of which are only present at specific stages during development. Mounting evidence indicates that alternate transcriptional pathways exist that are driven by ecdysteroids other than 20E. Coordinate changes in ecdysteroid-regulated gene expression occur at several stages in the *Drosophila* life cycle at times when 20E titer is known to be low. Given the reported biological activity of these ecdysteroids, it seems reasonable to expect that some of the diverse pathways can be mediated by a single EcR molecule. Due to the presence of the pliable ligand- and DNA binding domains EcR could adopt different, though ligand and response element specific, conformations and thus in cooperation with specific cofactors and/or primary transcription factors act as a universal and versatile factor controlling numerous ecdysteroid-dependent genes in a tissue- and gene specific pattern.

## MATERIALS AND METHODS

### *Bacterial strain and plasmid vector*

The plasmid pGEX-2T (Amersham Biosciences, Germany) containing $lacI^q$ gene was used for the expression of DBDs as fusions with *Schistosoma japonicum* glutathione S-transferase (GST). For the DBDs production *Escherichia coli* strain BL21(DE3)pLysS (Novagen, Germany) was used.

*Construction of DBDs-expresion vectors; Site-directed mutagenesis*

The construction of the expression plasmids for the wild-type *Drosophila melanogaster* EcR and Usp GST-DBDs (pGEX-2T-EcRDBD, pGEX-2T-UspDBD respectively) has been described previously. The expression plasmid for the wild-type *Bombyx mori* GST-DBD was constructed analogously using following primers 5'-gcccgggg*atcc*GCACCTCGA CAGCAAGAG-3'(sense) and 5'-gcccgggg*atcc*GTCTTCGACTGTGGTCGTA-3' (antisense). The construction of the expression plasmid for the 6×His-tagged DBD (amino acids 212 do 337) of the *Drosophila* orphan nuclear receptor DHR38 (DHR38DBD)will be described elsewhere.

PCR-based megaprimer mutagenesis protocol was used to introduce new codons within the DNA region encoding the putative T-box and the recognition α-helix of the EcRDBD. Plasmid pGEX-2T–EcRDBD was used as a template. The sequence of the recombinant DNA fragments was verified by dideoxy sequencing.

*Overexpression and purification of the wild-type and mutant proteins*

The expression of GST-DBDs and purification of the wild-type *Drosophila* UspDBD, EcRDBD and its mutated derivatives, as GST-free proteins, were performed as described previously. The *Bombyx mori* EcRDBD was isolated using procedure elaborated for the *Drosophila* EcRDBD. The procedure for the overexpression and purification of the DHR38DBD will be described elsewhere.

*DNA-binding assays*

Electrophoretic mobility-shift assays (EMSAs) were performed under conditions described previously. The indicated amounts of protein(s) and constant amount (fmol) of the appropriate $^{32}$P-labelled double-stranded oligonucleotide and 110 ng of poly(dI-dC) (Amersham Biosciences, Germany) were incubated for 30 min on ice in a final volume of 25

μL binding buffer (50 mM $Na_2HPO_4$, 100 mM NaCl, 1 mM 2-mercaptoethanol, 5 μM $ZnCl_2$, 10% (v/v) glicerol, pH 7.8 at 22°C). The DBD-DNA complexes were separated from the free DNA on 5% polyacrylamide gel in 0.25 x Tris/Borate/EDTA. The gel was precooled to 4°C overnight, prerun at 160 V for 20 min and after applying samples the electrophoresis was continued at 4°C for 30 min at 270 V and then for 240 min at 200 V. After that the gels were dried under vacuum at 80°C and exposed to Imaging Plates (Fuji Photo Film, Japan). Fluorescence signals were scanned with the Fuji Film FLA-3000 Fluorescent Image Analyzer (Raytest Isotopenme_geräte GmbH, Germany). Scans were read at 50-_m resolution and 16-bit quantitative image accuracy, and then analyzed using AIDA Bio-Package software (Raytest Isotopenme_geräte GmbH, Germany).

*Circular dichroism spectroscopy*

CD spectra were recorded using Jasco J-715 spectrometer in 0.1 cm cuvettes thermostated at 20°C in 50 mM $Na_2HPO_4$ buffer (pH 7.8 at 22°C) containing 250 mM NaCl, 1 mM 2-mercaptoethanol, 5 μM $ZnCl_2$, 10% (v/v) glycerol. Spectra were recorded with response time of 1.0 sec and with data spacing 0.5 nm. Each spectrum shown is the result of five spectra accumulated and averaged.

*Chemical denaturation*

Denaturation profiles were monitored by florescence ($\lambda_{ex}$ = 275 nm, $\lambda_{em}$ = 305 nm, at 18°C) using the Fluorolog-3 instrument (SPX, Jobin-Yvon, Horiba, ISA) and quartz cuvettes with 1 cm path length. The concentration of the proteins was 11.5 μM (9.0 μM for V72I/P73Q/N75P mutant). The concentrated guanidinium hydrochloride solution (7.0 M) was added to the protein samples in the phosphate buffer (50 mM $Na_2HPO_4$, 250 mM NaCl, 1 mM 2-mercaptoethanol, 5 μM $ZnCl_2$, 10% (v/v) glycerol, pH 7.8 at 22°C). To obtain desired denaturant concentration the defined volumes of the samples were withdrawn from the

incubation mixture and corresponding volumes of the guanidinium hydrochloride solution were added to the mixture to maintain final volume 500 μL. Each data point was calculated considering changes in protein and denaturant concentration.

*Protein concentration*

Protein concentration of the purified proteins was determined spectrophotometrically at 280 nm using absorption coefficients calculated according to Gill and von Hippel.

ACKNOWLEDGEMENTS

We thank Luc Swevers and Kostas Iatrou (Institute of Biology, National Center for Scientific Research "Demokritos" 153 10 Aghia Paraskevi Attikis, Athens, Greece) for the cDNA clone encoding the full length of *Bombyx mori* EcR (pBluescript-$SK^+$_BmEcR).

This study was supported by grant (3P04B 009 23) from the Polish State Committee for Scientific Research.

REFERENCES

(.....)

FIGURE LEGENDS

*Fig. 1. Schematic diagram of the macromolecular components used in this study.*

**(A)** Amino acid sequence of the wild-type *Drosophila melanogaster* EcRDBD. The numbering is relative to the first $Zn^{2+}$-coordinating cysteine of the DBD. Residues in the open circles correspond to the P-box amino acids. Residues from the DNA recognition α-helix (boxed) and residues from T-box region that where substituted with alanine arein bold. Five amino acids were excluded from the alanine scanning. In particular, C18, and C21 which are two of the eight absolutely conserved cysteines which coordinate two $Zn^{2+}$ ions in nuclear receptors DBDs, and highly conserved F25, F26, V29 residues that according to data published for other receptors, including our observations concerning UspDBD, contribute to

hydrophobic interaction that stabilize the DBD structure. (B) Sequences of the oligonucleotides probes used in EMSAs (only one strand is shown). The sequence of the $hsp27_{pal}$ oligonucleotide is based on the natural 20E pseudo-palindromic response element (marked with the arrows) from the *D. melanogaster hsp27* gene promoter. $hsp27_{\Delta R}$ is the derivative of the $hsp27_{pal}$ sequence which contains only left half-site of the $hsp27_{pal}$ (marked with the arrow).

*Fig. 2. DNA-binding activities of the mutants of the EcRDBD DNA recognition a-helix.*

Electrophoretic mobility shift assays were carried out with $hsp27_{pal}$ (A, B, E and F), $hsp27_{\Delta R}$ (C and D) and with the indicated homogenous EcRDBD (A, B, C and D) or equimolar mixture of the respective EcRDBD and the wild-type UspDBD (E and F). (B), (D) and (F) represent quantitative analysis of the EMSA data presented in (A) (lanes 2-11), (C) (lanes 2-12), and (E) ( lanes 2-11), carried out using Fuji Film FLA-3000 Fluorescent Image Analyzer. The columns indicate mean values of three independent experiments and bars indicate standard deviations.

The complexes formed by one DBD molecule are indicated by CI and those originated from homo- or heterodimer by CII; F, free probe. The DBDs forming the respective complex are denoted by U for UspDBD, E for EcRDBD. The protein concentrations were: (A) lanes 2-12, 240 nM of the indicated EcRDBD; lane 13, the same amount of the wild-type UspDBD; lane 14, 120 nM of each wild-type DBD; (C) protein concentrations were as in (A); (E) lanes 2-11, 60 nM of the wild-type UspDBD and 60 nM of the indicated EcRDBD; lane 12, 120 nM of the wild type UspDBD; lane 13, 120 nM of the wild-type EcRDBD.

Fig. 3. *Far-UV CD spectra of the mutants of the EcRDBD DNA recognition a-helix.*

CD measurements of protein solutions (10 μM) were done on the Jasco J-715 spectropolarimeter. Spectra were recorded with the response time of 1.0 sec and with the data point resolution of 1 nm using cuvette with 0.1 cm path length. Five scans were averaged in

order to obtain smooth spectra. The black circles represent the spectra of the wild-type EcRDBD, the open circles represent the spectra of the indicated mutant. $[\theta]_{MRE}$ indicates mean residue ellipticity (in deg_cm$^2$_dmol$^{-1}$).

*Fig. 4. DNA-binding of the EcRDBD T- box mutants to hsp27$_{pal}$.*

Gel retardation experiments were performed with *hsp27$_{pal}$* and increasing amounts of the indicated EcRDBD (A-H). The respective complexes were visualized by phosphoimaging (insets) and their quantitaion is graphically depicted. Black circles and black triangles represent CII and CI complexes formed by the wild-type EcRDBD; the open circles and the open triangles represent CII and CI complexes formed by the indicated mutant. Protein concentrations (in nM) in lanes 1-13 were: 0, 8, 16, 32, 60, 120, 200, 240, 400, 500, 600, 800, 1000. The figure shows a representative experiment; similar results have been obtained in at lest four independent experiments. For more details see text and the legend to Fig. 2.

*Fig.5. DNA-binding of the EcRDBD T-box mutants in the presence of the wild-type UspDBD.*

Gel retardation experiments were performed with *hsp27$_{pal}$* and increasing amounts of equimolar mixture of the indicated EcRDBD with the wild-type UspDBD. Protein concentration (in nM) were: lanes 1-13: 0, 4, 8, 16, 30, 60, 100, 120, 200, 250, 300, 400, 500 of each DBDs. The black circles represent the wild-type UspDBD/EcRDBD heterodimer; the open circles represent the complexes formed by the wild-type UspDBD with the indicated mutant. Other details as in the legend to Fig. 4.

*Fig. 6. Far-UV CD spectra of the EcRDBD T- box mutants.*

The black circles represent the spectra of the wild-type EcRDBD, the open circles represent the spectra of the indicated mutant, respectively. $[\theta]_{MRE}$ is mean residue ellipticity (in deg_cm$^2$_dmol$^{-1}$). Other details as in the legend to Fig. 3.

*Fig. 7. The triple mutant of the Drosophila EcRDBD.* The triple mutant (V72I/P73Q/N75P) of the *Drosophila* EcRDBD (*D. m.*) containing at positions 72, 73 and 75 amino acid residues

occurring in *Bombyx mori* (*B. m.*) EcRDBD (A) was obtained in a homogenous form (not shown). The $hsp27_{pal}$-binding affinity of the V72I/P73Q/N75P mutant (open circles)) was analyzed in the absence (B) and in the presence (C) of the wild-type UspDBD and compared with the affinity wild-type *Drosophila* EcRDBD (filled circles). Panel (D) shows CD spectra of the *Drosophila* EcRDBD and of the triple mutant. For more details see legends to Fig. 3 and 4.

*Fig. 8. Dissymmetry in the molecular properties of the EcRDBD and UspDBD.*

(A) Comparison of the EcRDBD and the UspDBD CD spectra. CD spectra of the wild-type *Drosophila* EcRDDBD (filled circles) and UspDBD (open circles) were recorded at 10 μM concentrations under standard conditions (see Materials and Methods). (B) Chemical denaturation profiles of the DBDs. Chemical denaturation of the wild-type and mutated DBDs was monitored by fluorescence emission measurements under conditions described in Materials and Methods section. The panel shows denaturation profiles for the following DBDs: wild-type *Drosophila* EcRDBD, filled circles; wild-type *Bombyx* EcRDBD, open triangles; triple mutant (V72I/P73Q/N75P ) of the *Drosophila* EcRDBD, filled triangles; *Drosophila* EcRDBD which does not contain A-box sequence, open circles; wild-type *Drosophila* UspDBD, filled squares. (C) Renaturation of DBDs. Electrophoretic mobility shift assays were performed with $hsp27_{pal}$ and with the indicated wild-type *Drosophila melanogaster* (*D. m.*) or *Bombyx mori* (*B.m.*) DBDs. The panel shows results obtained for the DBDs which were incubated at 8.7 μM concentration on ice for 1h with 4.0 M guanidinum hydrochloride and then diluted to obtain 0.1 M guanidinum hydrochloride concentration (denoted as renaturated DBD), or for the DBDs which were treated in the same manner except that guanidinum hydrochloride was completely omitted or present at 0.1 M concentration (denoted as 0.1 M GdmCl). The protein concentrations were: lanes 1-10, 218 nM of the indicated DBD; lanes 11- 14, 109 nM of the indicated DBD. (D) Quantitative analysis of the

data presented in (C).The quantitative analysis was carried out using Fuji Film FLA-3000 Fluorescent Image Analyzer. The columns indicate mean values of three independent experiments and bars indicate standard deviations.

Fig. 9. *Interaction of the EcRDBD with the DHR38DBD on hsp27pal.*

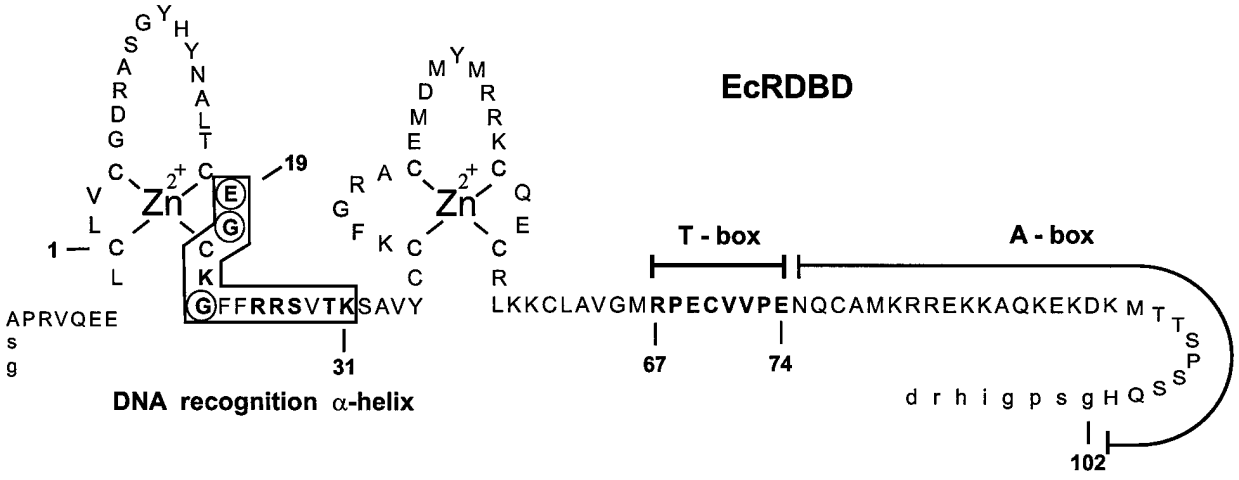
The $^{32}$P-labelled $hsp27_{pal}$ and 500 nM of the wild-type EcRDBD were incubated (lanes 3-8) with increasing amounts (molar excess is shown over respective lanes) of the DHR38DBD. The complexes formed by one DBD molecule are indicated by CI and those originated from homo- by CII and heterodimer by CIII; F, free probe. The DBDs forming the respective complexes are denoted by E for EcRDBD, D for DHR38DBD. The protein concentrations were: lane 3, 500 nM of both DBDs; lane 9, 500 nM of DHR38DBD; lane 10, 2000 nM of DHR38DBD, respectively.

Fig. 10. A*lignment of the C-terminal extension (CTE) sequences of the DBDs.*

The alignment was done using ClustalX(1.81) (ftp://ftp-igbmc.u-strasbg.fr/pub/ClustalX/). ' * ' indicates positions, which have a single, fully conserved residue; ' : ' indicates that one of the following 'strong' groups is fully conserved. Amino-acid residue numbers are relative to the first Zn-coordinating cysteine. Regions corresponding to to the so-called T-box are boxed. Sequences were taken from SWISS-PROT (http://www.expasy.org/sprot/). Abbreviations used: *Drosophila melanogaster* (Fruit fly) (Dm, accession number: EcR: P34021; Usp: P20153), *Bombyx mori* (Silk moth) (Bm, P49881), *Manduca sexta* (Tobacco hornworm) (Ms, P49883), *Aedes aegypti* (Yellowfever mosquito) (Ae, P49880), *Locusta migratoria* (Migratory locust) (Lm, O97095), *Tenebrio molitor* (Yellow mealworm), *Heliothis virescens* (Noctuid moth) (Hv, O18473), *Chilo suppressalis* (Chs, Q8MYA6), *Choristoneura fumiferana* (Spruce budworm) (Chf, O77240), *Calliphora vicina* (Blue blowfly) (Cv, Q9GPH1), *Ceratitis capitata* (Mediterranean fruit fly) (Cc, O76827), *Lucilia cuprina* (Greenbottle fly) (Lc, O18531), *Chironomus tentans* (Midge) (Cht, P49882), *Aedes albopictus*

(Forest day mosquito) (Aal, Q9U3Y4), *Amblyomma americanum* (L.) (Ixodid tick), (Aam, 044338), human retinoic acid receptor-alpha (hRXR, P19793), human thyroid hormone receptor beta-1(hTR, P10828), human farnesoid X-activated receptor (hFXR_β, Q96RI1-2|Q96RI1 splice isoform 2 of Q96RI1), human vitamin D3 receptor (hVDR, P11473), probable nuclear hormone receptor HR38 from *Drosophila melanogaster* (Fruit fly) (DHR38, P49869).

Ⓐ


EcRDBD
T - box
A - box
DNA recognition α-helix
APRVQEE
s
g
FFRRSVTKSAVY
LKKCLAVGMRPECVVPENQCAMKRREKKAQKEKDK
1
19
31
67
74
102

Ⓑ

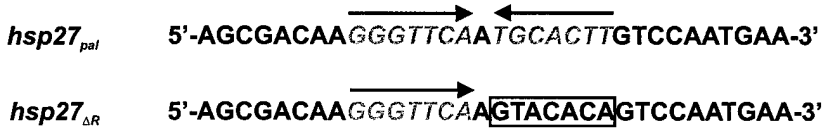

hsp27pal 5'-AGCGACAAGGGTTCAATGCACTTGTCCAATGAA-3'
hsp27ΔR 5'-AGCGACAAGGGTTCAAGTACACAGTCCAATGAA-3'

FIGURE 1

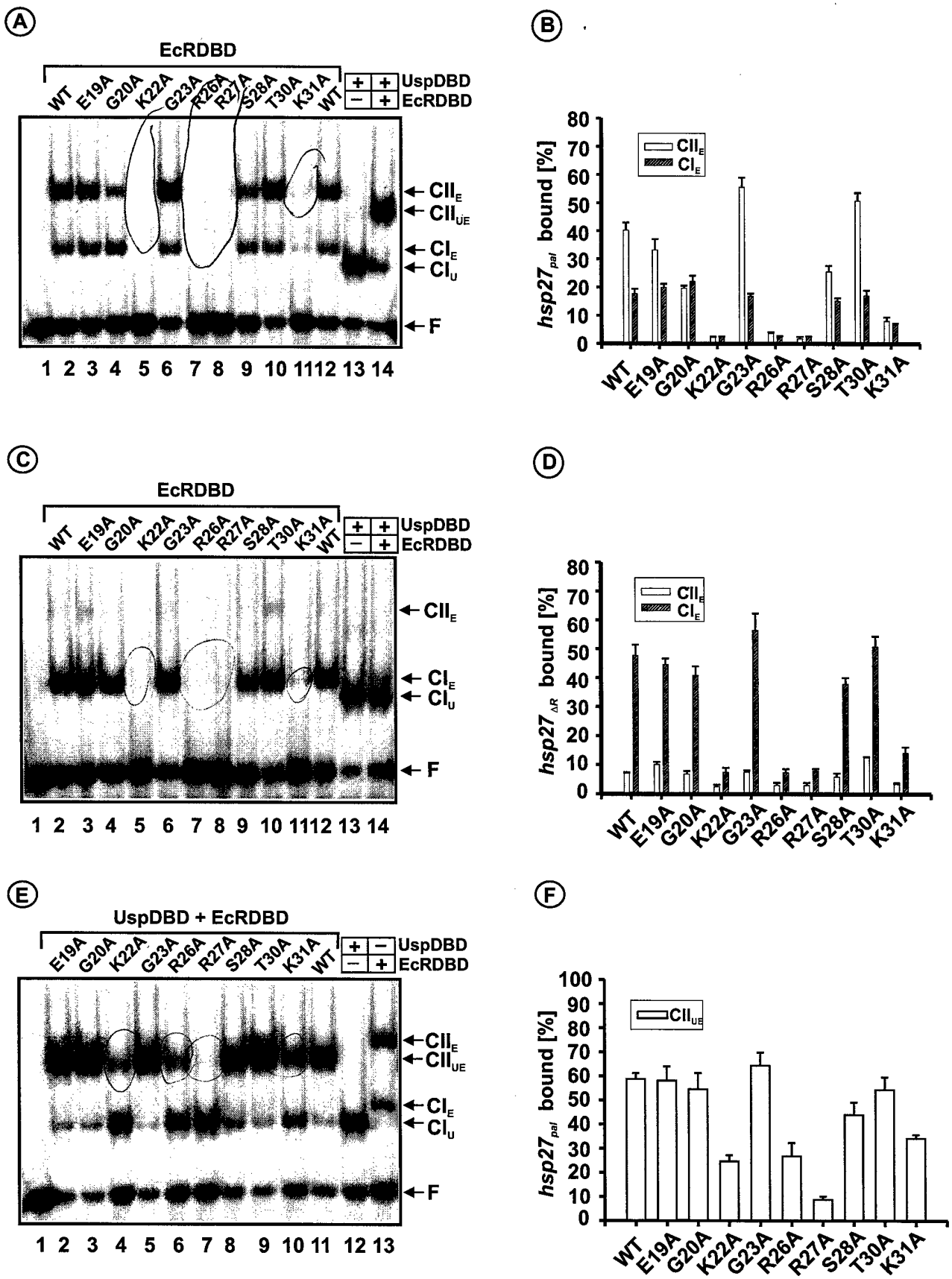

A
EcRDBD
WT E19A G20A K22A G23A R26A R27A S28A T30A K31A WT
+ + UspDBD
− + EcRDBD
CII_E
CII_UE
CI_E
CI_U
F
1 2 3 4 5 6 7 8 9 10 1112 13 14
B
hsp27_pal bound [%]
CII_E
CI_E
WT E19A G20A K22A G23A R26A R27A S28A T30A K31A
C
EcRDBD
WT E19A G20A K22A G23A R26A R27A S28A T30A K31A WT
+ + UspDBD
− + EcRDBD
CII_E
CI_E
CI_U
F
1 2 3 4 5 6 7 8 9 10 1112 13 14
D
hsp27_ΔR bound [%]
CII_E
CI_E
WT E19A G20A K22A G23A R26A R27A S28A T30A K31A
E
UspDBD + EcRDBD
E19A G20A K22A G23A R26A R27A S28A T30A K31A WT
+ − UspDBD
− + EcRDBD
CII_E
CII_UE
CI_E
CI_U
F
1 2 3 4 5 6 7 8 9 10 11 12 13
F
hsp27_pal bound [%]
CII_UE
WT E19A G20A K22A G23A R26A R27A S28A T30A K31A

FIGURE 2

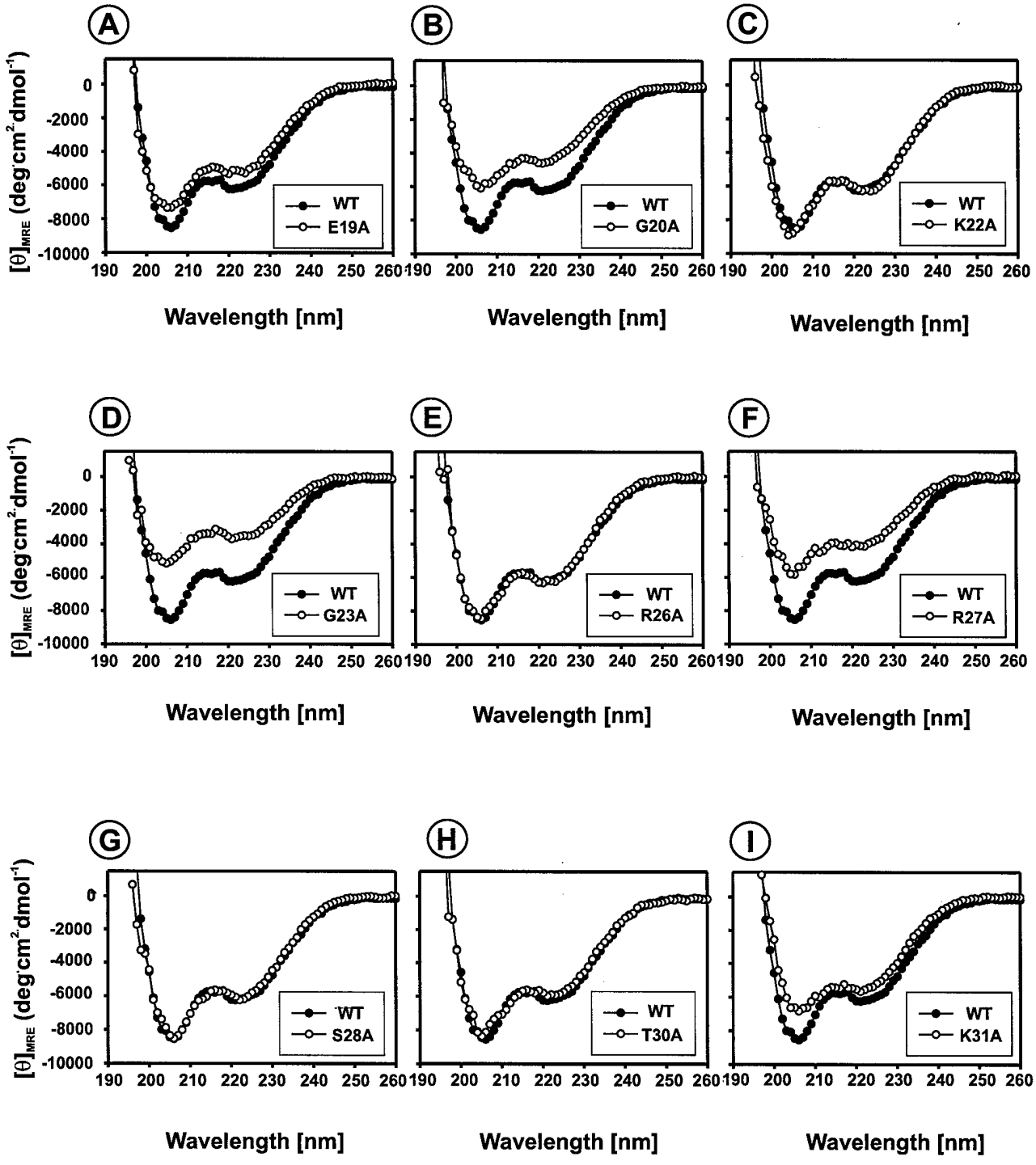

A
B
C
D
E
F
G
H
I
$[\theta]_{MRE}$ (deg·cm$^2$·dmol$^{-1}$)
0
-2000
-4000
-6000
-8000
-10000
190 200 210 220 230 240 250 260
Wavelength [nm]
WT
E19A
G20A
K22A
G23A
R26A
R27A
S28A
T30A
K31A

FIGURE 3

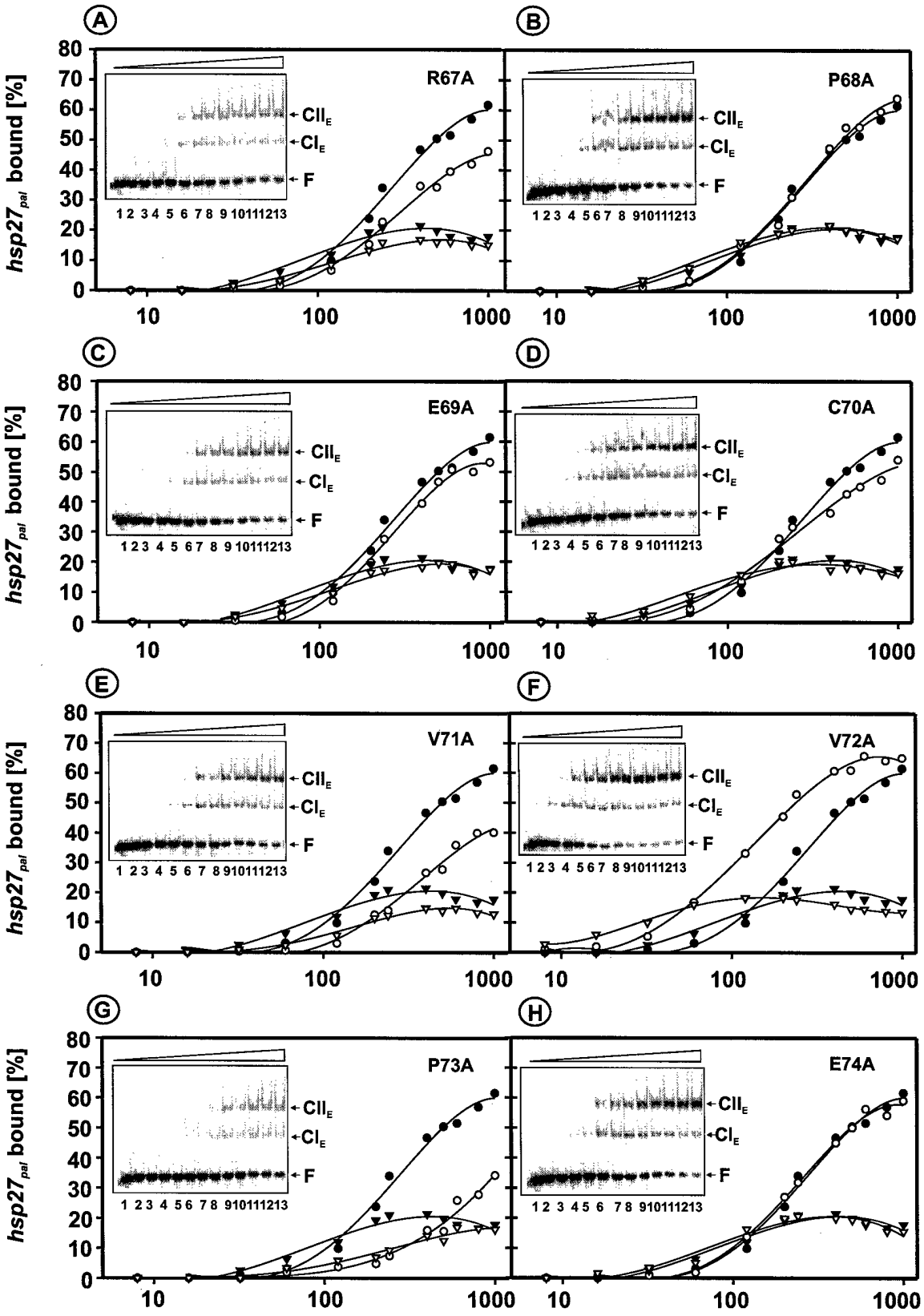

A
B
C
D
E
F
G
H
R67A
P68A
E69A
C70A
V71A
V72A
P73A
E74A
CII_E
CI_E
F
1 2 3 4 5 6 7 8 9 10 11 12 13
hsp27_pal bound [%]
0
10
20
30
40
50
60
70
80
10
100
1000
EcRDBD concentration [nM]

FIGURE 4

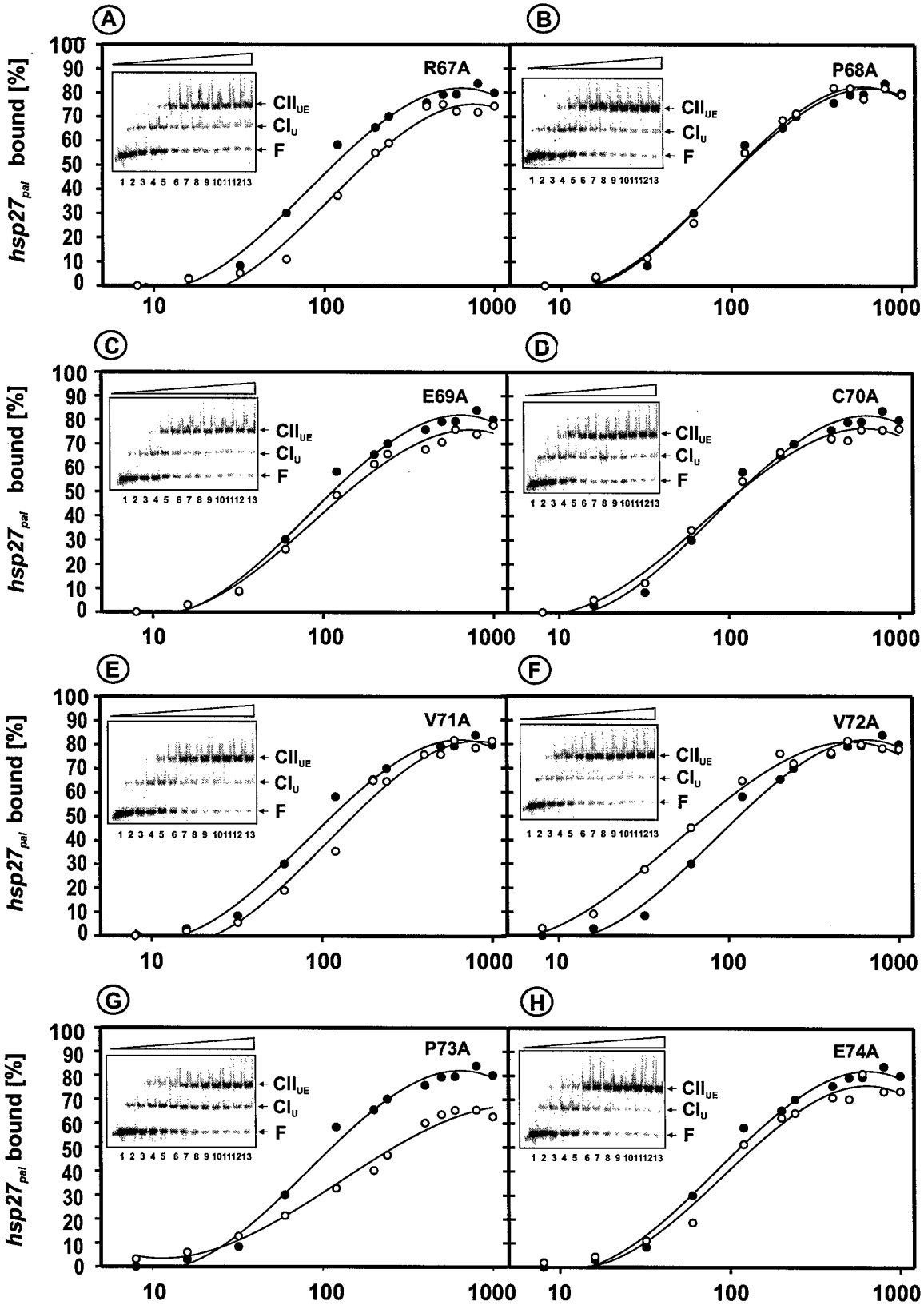

A
B
C
D
E
F
G
H
R67A
P68A
E69A
C70A
V71A
V72A
P73A
E74A
$CII_{UE}$
$CI_U$
F
1 2 3 4 5 6 7 8 9 10 11 12 13
$hsp27_{pal}$ bound [%]
0
10
20
30
40
50
60
70
80
90
100
10
100
1000
UspDBD/EcRDBD concentration [nM]

FIGURE 5

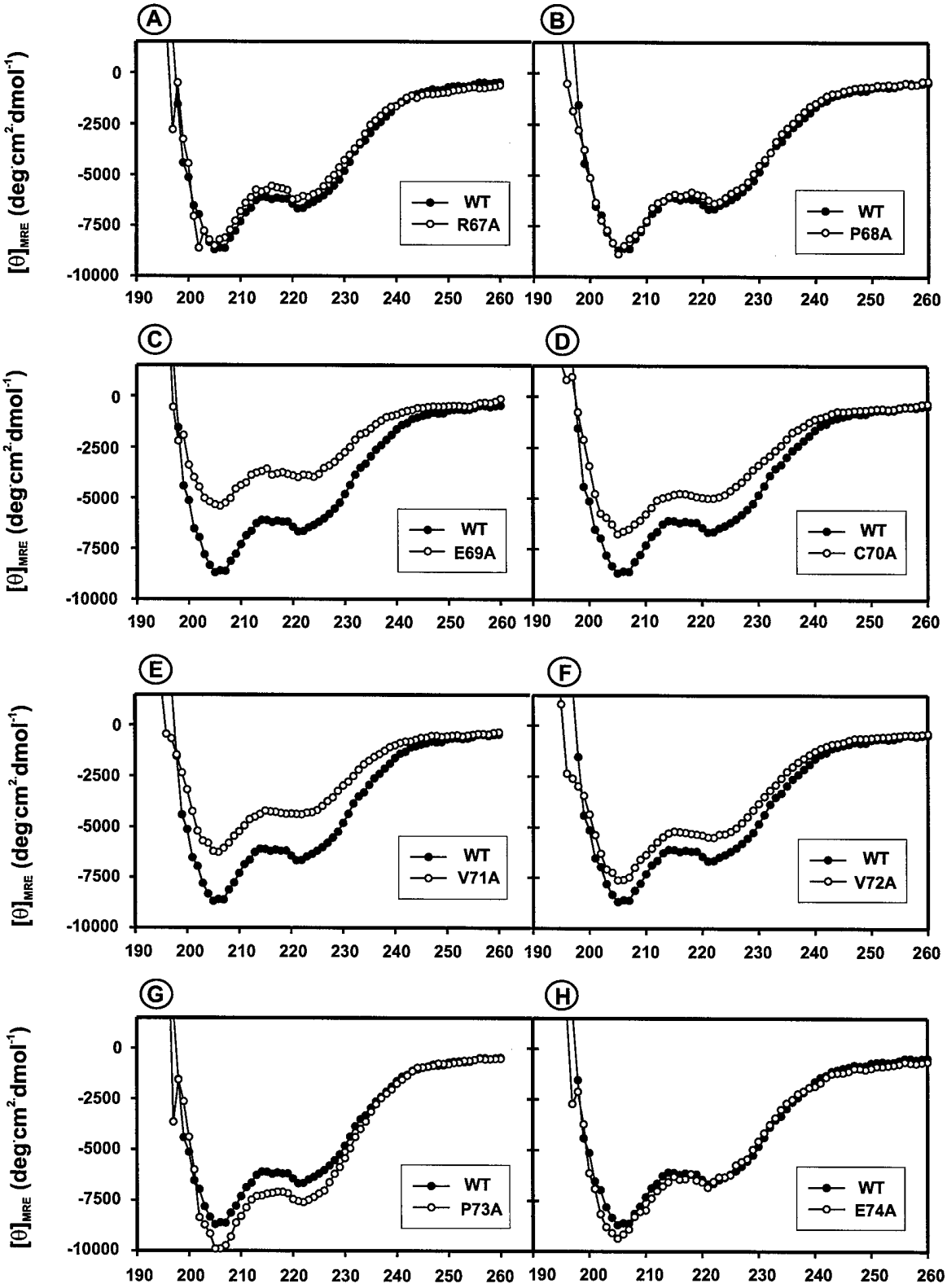
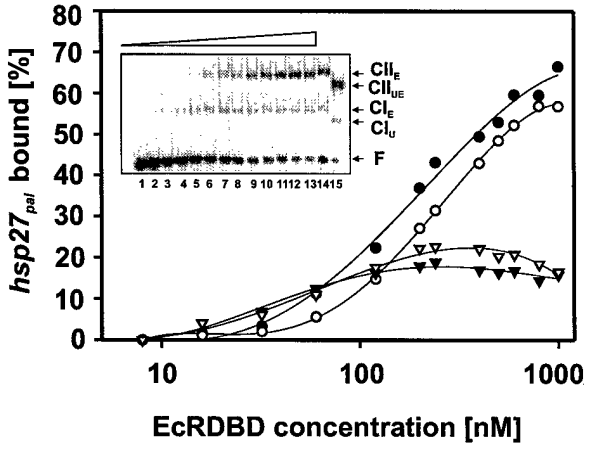

A
B
C
D
E
F
G
H
WT
R67A
P68A
E69A
C70A
V71A
V72A
P73A
E74A
0
-2500
-5000
-7500
-10000
190
200
210
220
230
240
250
260
[θ]MRE (deg·cm2·dmol-1)
Wavelength [nm]

FIGURE 6

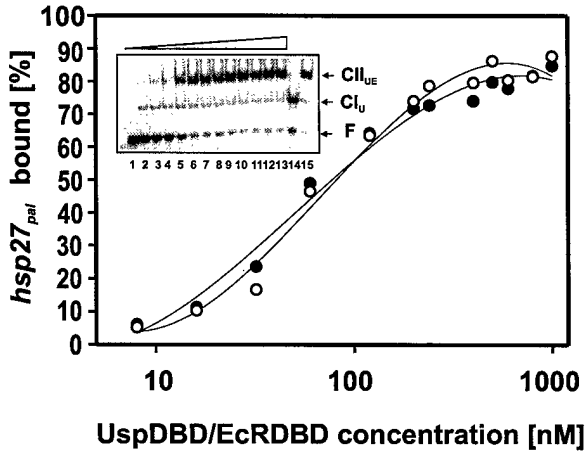
Ⓐ

| EcRDBD | Sequences |
|---|---|
| | T - box A-box |
| *D.m.* | CRLKKCLAVGMRPECV*VPE*_N_QCAMKRREKKAQKEKDKMTTSPSSQH |
| | 67 77 87 97 |
| *B.m.* | CRLKKCLAVGMRPECVIQEPSKNKDRQRQKKDKGILLPVSTTTVED |
| V72I/P73Q/N75P | CRLKKCLAVGMRPECV*IQE*_P_QCAMKRREKKAQKEKDKMTTSPSSQH |

Ⓑ




Ⓒ




Ⓓ

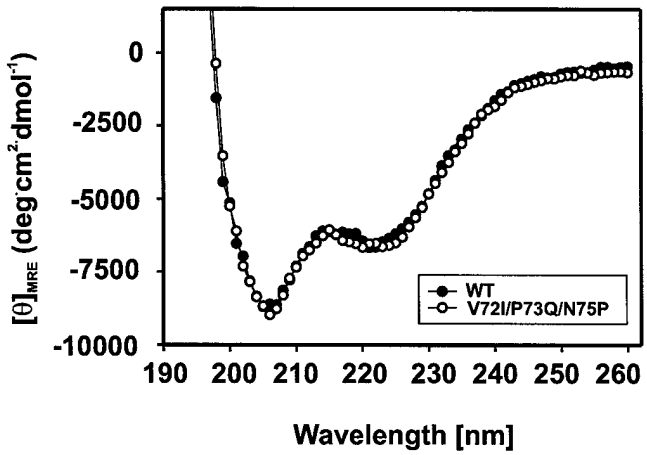



FIGURE 7

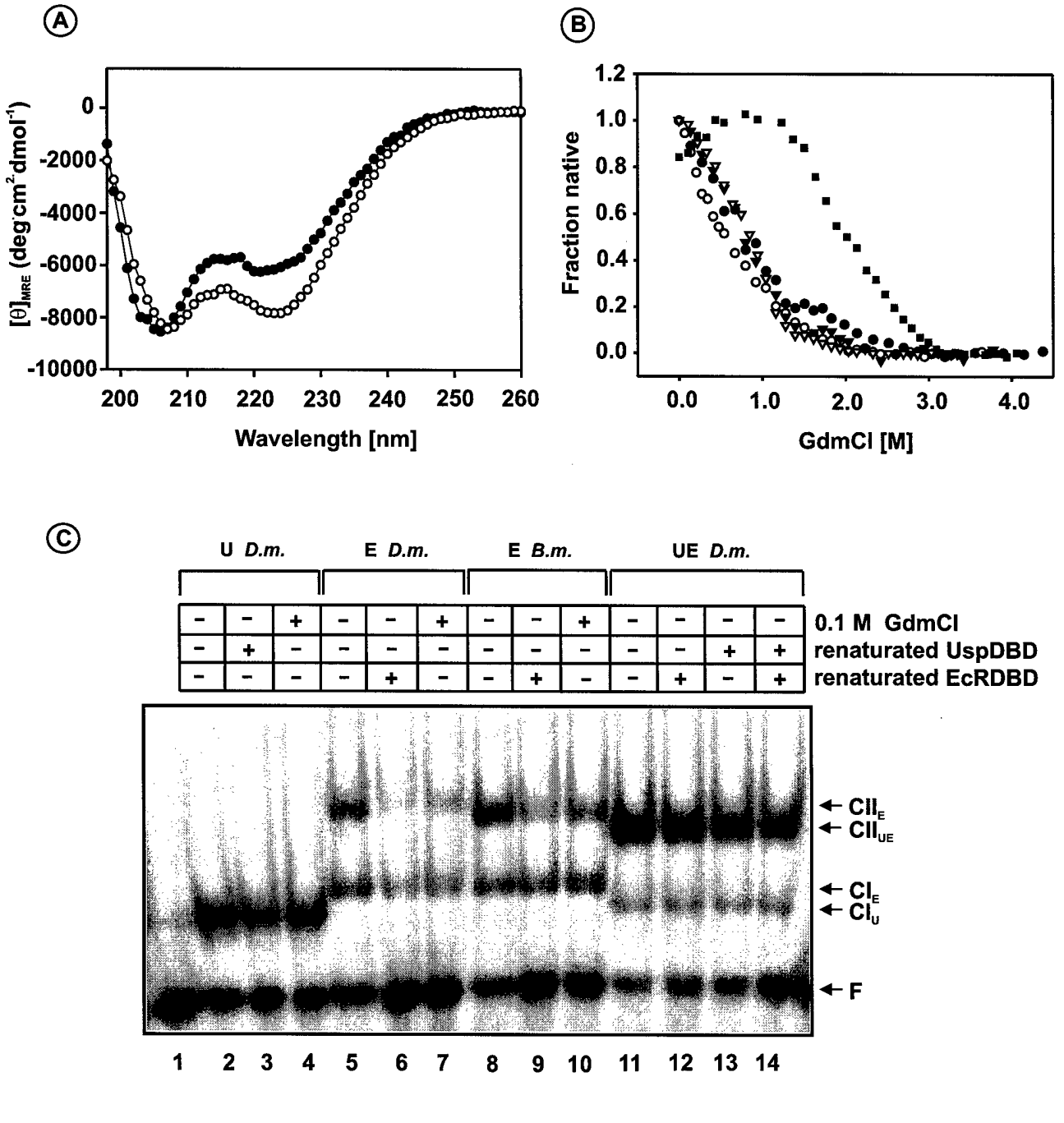

A
B
C
D
$[\theta]_{MRE}$ (deg·cm$^2$·dmol$^{-1}$)
Wavelength [nm]
Fraction native
GdmCl [M]
U D.m.
E D.m.
E B.m.
UE D.m.
0.1 M GdmCl
renaturated UspDBD
renaturated EcRDBD
$CII_E$
$CII_{UE}$
$CI_E$
$CI_U$
F
hsp27$_{pal}$ bound [%]

FIGURE 8

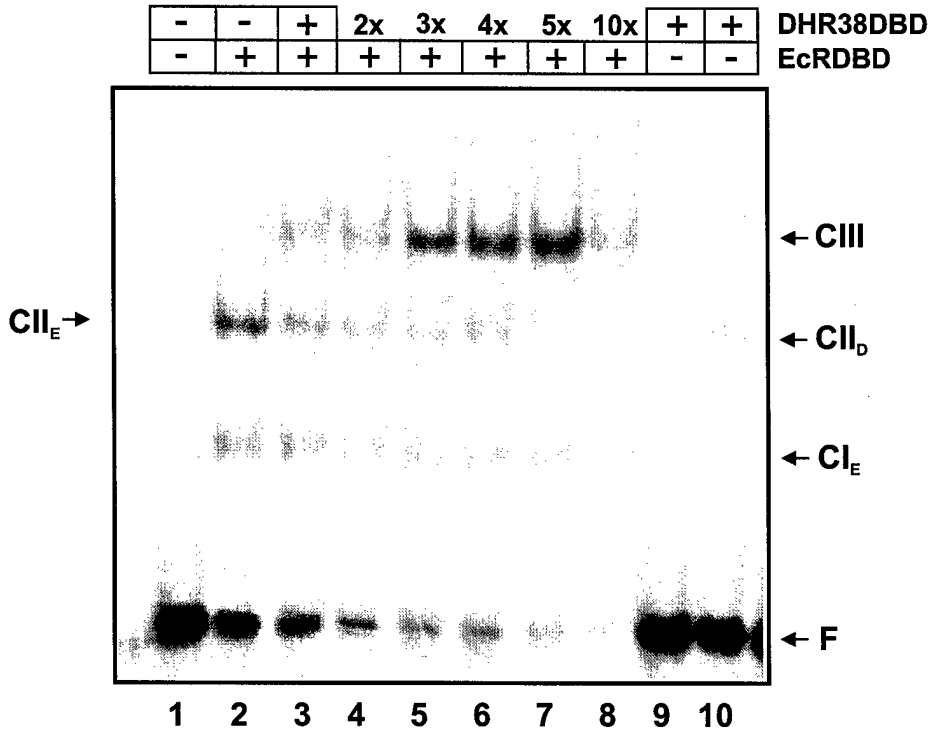

- - + 2x 3x 4x 5x 10x + +
DHR38DBD
- + + + + + + + - -
EcRDBD
CII$_E$
CIII
CII$_D$
CI$_E$
F
1 2 3 4 5 6 7 8 9 10

FIGURE 9

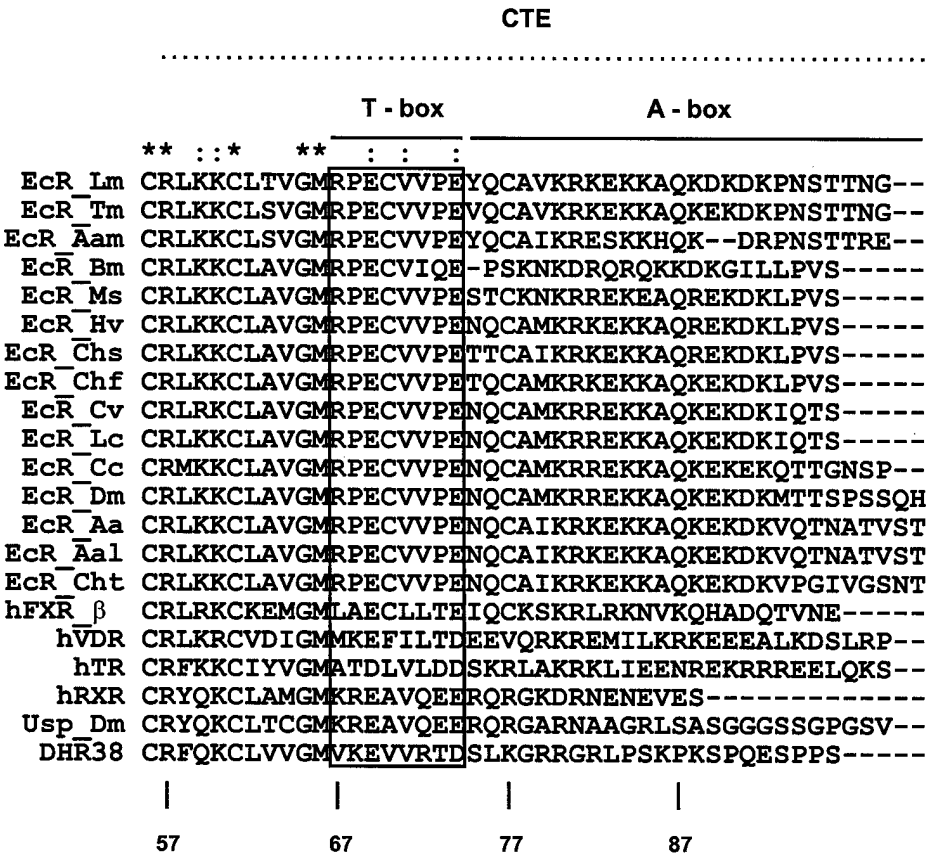

CTE
T - box
A - box
EcR_Lm CRLKKCLTVGMRPECVVPEYQCAVKRKEKKAQKDKDKPNSTTNG--
EcR_Tm CRLKKCLSVGMRPECVVPEVQCAVKRKEKKAQKEKDKPNSTTNG--
EcR_Aam CRLKKCLSVGMRPECVVPEYQCAIKRESKKHQK--DRPNSTTRE--
EcR_Bm CRLKKCLAVGMRPECVIQE-PSKNKDRQRQKKDKGILLPVS-----
EcR_Ms CRLKKCLAVGMRPECVVPESTCKNKRREKEAQREKDKLPVS-----
EcR_Hv CRLKKCLAVGMRPECVVPENQCAMKRKEKKAQREKDKLPVS-----
EcR_Chs CRLKKCLAVGMRPECVVPETTCAIKRKEKKAQREKDKLPVS-----
EcR_Chf CRLKKCLAVGMRPECVVPETQCAMKRKEKKAQKEKDKLPVS-----
EcR_Cv CRLRKCLAVGMRPECVVPENQCAMKRREKKAQKEKDKIQTS-----
EcR_Lc CRLKKCLAVGMRPECVVPENQCAMKRREKKAQKEKDKIQTS-----
EcR_Cc CRMKKCLAVGMRPECVVPENQCAMKRREKKAQKEKEKQTTGNSP--
EcR_Dm CRLKKCLAVGMRPECVVPENQCAMKRREKKAQKEKDKMTTSPSSQH
EcR_Aa CRLKKCLAVGMRPECVVPENQCAIKRKEKKAQKEKDKVQTNATVST
EcR_Aal CRLKKCLAVGMRPECVVPENQCAIKRKEKKAQKEKDKVQTNATVST
EcR_Cht CRLKKCLAVGMRPECVVPENQCAIKRKEKKAQKEKDKVPGIVGSNT
hFXR_β CRLRKCKEMGMLAECLLTEIQCKSKRLRKNVKQHADQTVNE-----
hVDR CRLKRCVDIGMMKEFILTDEEVQRKREMILKRKEEEALKDSLRP--
hTR CRFKKCIYVGMATDLVLDDSKRLAKRKLIEENREKRRREELQKS--
hRXR CRYQKCLAMGMKREAVQEERQRGKDRNENEVES----------
Usp_Dm CRYQKCLTCGMKREAVQEERQRGARNAAGRLSASGGGSSGPGSV--
DHR38 CRFQKCLVVGMVKEVVRTDSLKGRRGRLPSKPKSPQESPPS-----
57
67
77
87

FIGURE 10

The EMBO Journal Vol. 22 No. 21 pp. 5827–5840, 2003

APPENDIX 5

# Structure of the heterodimeric ecdysone receptor DNA-binding complex

**Srikripa Devarakonda[1], Joel M.Harp[2], Youngchang Kim[3], Andrzej Ożyhar[4] and Fraydoon Rastinejad[1,2,5]**

[1]Department of Pharmacology, [2]Department of Molecular Genetics and Biochemistry, University of Virginia Health System, Charlottesville, VA 22908, [3]Argonne National Laboratory, Biosciences Division/Structural Biology Center, Argonne, IL 60439, USA and [4]Institute of Organic Chemistry, Biochemistry and Biotechnology, Division of Biochemistry, Wroclaw University of Technology, 50-370 Wroclaw, Poland

[5]Corresponding author
e-mail: fr9c@virginia.edu

**Ecdysteroids initiate molting and metamorphosis in insects via a heterodimeric receptor consisting of the ecdysone receptor (EcR) and ultraspiracle (USP). The EcR–USP heterodimer preferentially mediates transcription through highly degenerate pseudo-palindromic response elements, resembling inverted repeats of 5′-AGGTCA-3′ separated by 1 bp (IR-1). The requirement for a heterodimeric arrangement of EcR–USP subunits to bind to a symmetric DNA is unusual within the nuclear receptor superfamily. We describe the 2.24 Å structure of the EcR–USP DNA-binding domain (DBD) heterodimer bound to an idealized IR-1 element. EcR and USP use similar surfaces, and rely on the deformed minor groove of the DNA to establish protein–protein contacts. As retinoid X receptor (RXR) is the mammalian homolog of USP, we also solved the 2.60 Å crystal structure of the EcR–RXR DBD heterodimer on IR-1 and found the dimerization and DNA-binding interfaces to be the same as in the EcR–USP complex. Sequence alignments indicate that the EcR–RXR heterodimer is an important model for understanding how the FXR–RXR heterodimer binds to IR-1 sites.**

*Keywords*: ecdysone/EcR/nuclear receptor/RXR/USP

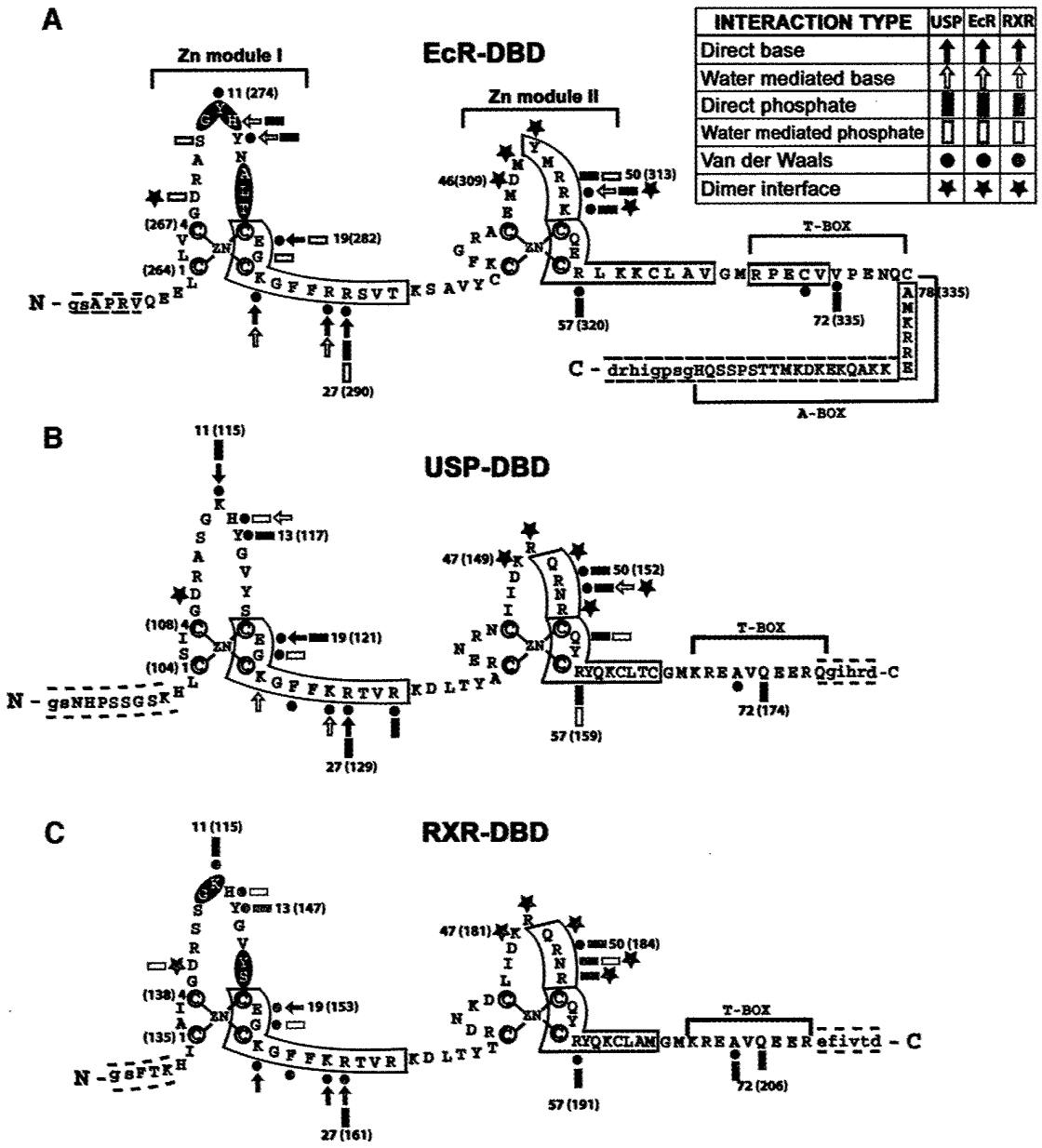
## Introduction

Ecdysteroids are arthropod-specific hormones that function as the major inducing and coordinating signals responsible for the widespread changes in gene expression associated with molting and metamorphosis in insects. In *Drosophila*, the ring gland secretes two major ecdysteroids, α-ecdysone and 20-deoxymakisterone A, which are largely inactive (Riddiford, 1996; Gilbert *et al.*, 1997). 20-Hydroxyecdysone (20E), which is the product of the α-ecdysone conversion in the peripheral tissues (Gilbert *et al.*, 2002), is thought to be a major signal for the coordinated programming of gene expression patterns responsible for the complete metamorphosis of the organism from larva to fly (Ashburner, 1973, 1974). The study of 20E-controlled gene activities has been pivotal in formulating modern concepts of steroid hormone-controlled gene activities and has functioned as a paradigm for steroid hormone activity (Riddiford, 1993, 1996). Signaling by 20E in *Drosophila* is mediated by the ecdysone receptor complex, a heterodimer of the ecdysone receptor (EcR; NR1H1) and ultraspiracle (USP; NR2B4) proteins, both of which are members of the nuclear receptor family (Koelle *et al.*, 1991; Yao *et al.*, 1992, 1993; Thomas *et al.*, 1993; Grad *et al.*, 2001). Although USP has no known ligands, it is the *Drosophila* homolog of the mammalian retinoid X receptors (RXRs) (Oro *et al.*, 1990), important members of the nuclear receptor family that bind to 9-*cis* retinoic acid and form heterodimeric complexes with other nuclear receptors, including the thyroid hormone receptor (TR), vitamin D3 receptor (VDR) and the all-*trans* retinoic acid receptor (RAR) (Mangelsdorf *et al.*, 1995).

Because there are only two types of consensus half-sites (5′-AGGTCA-3′ and 5′-AGAACA-3′) used by essentially all nuclear receptors (Glass, 1994), target selectivity appears to rely on the geometry and spacing of the half-sites, and not just the half-site sequences. Importantly, the DNA-binding domains (DBDs) typically generate the same pattern of DNA selectivity and subunit dimerization as the full-length receptors (Ożyhar *et al.*, 1991; Ożyhar and Pongs, 1993; Mader *et al.*, 1993; Perlmann *et al.*, 1993; Zechel *et al.*, 1994b, Grad *et al.*, 2001). There are exceptions to this rule, such as the VDR, whose DBD does not generate the same pattern of subunit dimerization as the intact receptor (Shaffer and Gewirth, 2002).

The RXR heterodimers generally target direct repeat DNA sites, which differ in their inter-half-site spacing, described by a 1–5 rule (Umesono and Evans, 1989; Umesono *et al.*, 1991; Mangelsdorf and Evans, 1995). A set of recent crystal structures of receptor DBDs on direct repeats has revealed how RXR and other non-steroid receptors bind to their cognate sites selectively. Specifically, the structures of RXR–RXR/DR-1, RXR–RAR/DR-1, RevErb–RevErb/DR-2, VDR–VDR/DR-3 and RXR–TR/DR-4 have been determined thus far, in each case showing DBDs forming non-symmetric 'head to tail' interactions (Rastinejad *et al.*, 1995, 2000; Zhao *et al.*, 1998, 2000; Sierk *et al.*, 2001; Shaffer and Gewirth, 2002).

In contrast, vertebrate steroid receptors, such as the glucocorticoid receptor (GR), the estrogen receptor (ER), the progesterone receptor (PR) and the mineralocorticoid receptor (MR), each form homodimers on their response elements. These response elements are inverted repeats with the same 3 bp (IR-3), causing their DBDs to form symmetric 'head to head' interactions (Luisi *et al.*, 1991; Schwabe *et al.*, 1993). Natural ecdysone response elements (EcREs) are imperfect inverted repeats of AGGTCA sequences separated by 1 bp (IR-1) (Riddihough and Pelham, 1987; Cherbas *et al.*, 1991; Ożyhar *et al.*, 1991;




D'Avino *et al.*, 1995; Fisk and Thummel, 1995; Horner *et al.*, 1995; Lehmann *et al.*, 1995, 1997; Hall and Thummel, 1998; Vögtli *et al.*, 1998). Therefore, the requirements for DNA binding for EcR–USP are distinct from those associated with both the RXR heterodimers and the steroid receptor homodimers. Besides the EcR–USP, only the RXR–FXR heterodimer is known to target IR-1 response elements preferentially (Laffitte *et al.*, 2000), where FXR is the nuclear receptor that binds directly to bile acids (Chawla *et al.*, 2001; Mi *et al.*, 2003).

Here, we describe the co-crystal structure of EcR–USP/IR-1 at 2.24 Å resolution. Importantly, RXR is the mammalian homolog of USP (Oro *et al.*, 1990), and these two receptors can also substitute for each other in stimulating high affinity DNA binding of EcR to IR-1 *in vitro* (Yao *et al.*, 1992, 1993; Thomas *et al.*, 1993). In addition, RXR and USP have both been shown to support ecdysone-responsive trans-activation equally in transfection assays in mammalian cell lines (Vögtli *et al.*, 1998). Moreover, USP has been shown to be capable of substituting for RXR in forming heterodimers with TR, RAR and VDR (Yao *et al.*, 1992; Thomas *et al.*, 1993; No *et al.*, 1996). Therefore, we also studied the 2.60 Å crystal structure of the EcR–RXR/IR-1 complex, to see to what extent the protein–protein and protein–DNA interactions were related to those in the EcR–USP/IR-1 complex. The structures together show how the DBDs form their protein–protein interactions in a manner that is highly discriminatory for the 1 bp spacing and inverted repeat geometry of the IR-1 site. Importantly, the dimerization surfaces used to engage EcR on DNA are essentially the same in USP and RXR but distinct from those observed in other RXR heterodimers or steroid receptor homodimers. Because FXR is closely related in sequence to EcR along

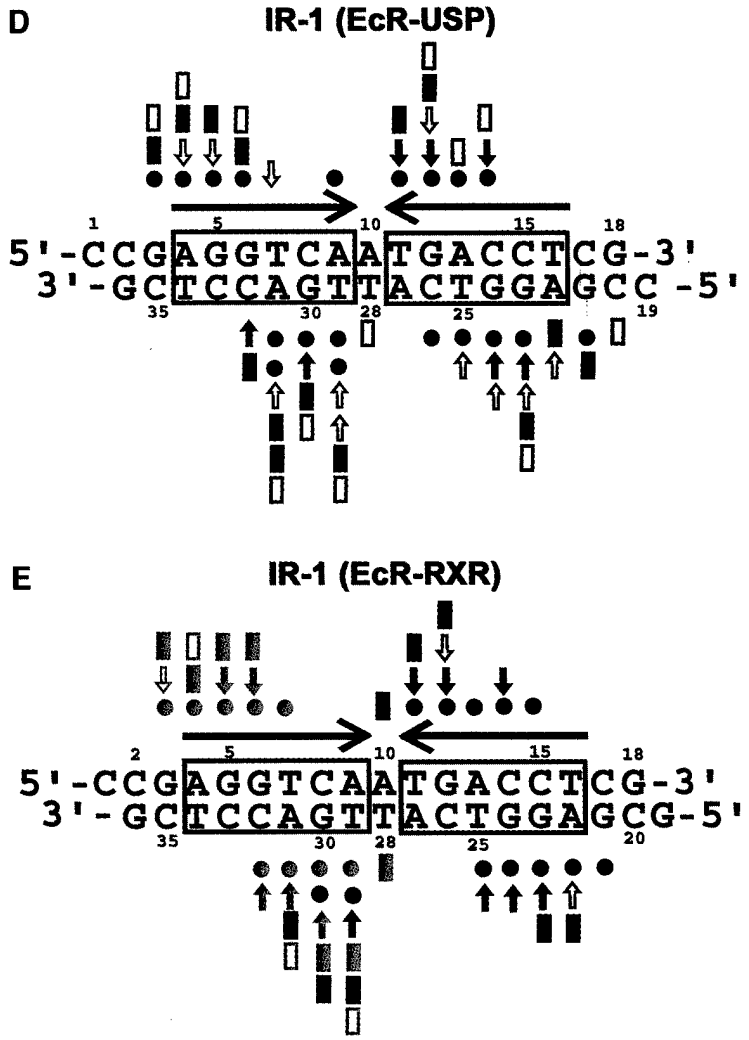



**Fig. 1.** The protein and DNA constructs used in crystallization and their contacts. (**A**) *Drosophila* EcR-DBD (blue), (**B**) *Drosophila* USP-DBD (red) and (**C**) human RXRα-DBD (gold) protein sequences. The yellow colored cysteines are those used for zinc coordination. The α-helices are boxed and the β-sheets are shaded. The numbering is relative to the first conserved cysteine, with the authentic numbers appearing in parentheses (Mangelsdorf *et al.*, 1995; Oro *et al.*, 1990; Koelle *et al.*, 1991). A key to the different interactions is provided within the figure. Dashed lines indicate N- and C-terminal residues not visible in the electron density maps. Cloning artifacts from the expression vector are indicated by lower case letters. The red boxed sequences in (A) indicate the A-box helix seen in the EcR–RXR complex. (**D**) The idealized IR-1 response element used in the EcR–USP complex, and (**E**) in the EcR–RXR complex. The symbols used are the same as in (A–C). Red symbols indicate contacts derived from USP, blue symbols indicate those from EcR, and gold indicates those from RXR.

its DBD and also binds as an RXR heterodimer to IR-1 (Laffitte *et al.*, 2000), our current structure of the EcR–RXR/IR-1 complex serves as a useful model for understanding how FXR and RXR are likely to cooperate in DNA binding.

## Results

### *Overall architecture of the EcR–USP and EcR–RXR complexes*

An 86 residue fragment of the *Drosophila* USP protein (residues 94–179) and a 109 residue fragment of *Drosophila* EcR (residues 256–364) were individually purified and used for co-crystallization on an IR-1 element (see Figure 1A, B and D, and Materials and methods). In the absence of EcR-DBD, USP-DBD can bind as a monomer to IR-1 (Niedziela-Majka *et al.*, 2000). Furthermore, in the absence of USP-DBD, EcR-DBD can bind IR-1, but primarily as homodimer (Niedziela-Majka *et al.*, 2001). However, when both DBDs are present in solution, the heterodimeric EcR–USP complex on IR-1 is formed with greater affinity and in a synergistic manner (Niedziela-Majka *et al.*, 2001). A number of natural EcREs have been identified to date, with most of these being imperfect palindromes of PuG(G/T)TCA with a single base pair acting as the spacer between the half-sites (Riddihough and Pelham, 1987; Cherbas *et al.*, 1991; Ożyhar *et al.*, 1991; Antoniewski *et al.*, 1993,1994, 1996; D'Avino *et al.*, 1995; Lehmann *et al.*, 1995, 1997; Vögtli *et al.*, 1998). Because idealized IR-1 elements form higher affinity EcREs, we carried out crystallization studies using the IR-1 sequence in which both half-sites were identical

**Table I.** Data collection and refinement statistics

| | EcR–USP | EcR–RXR | | |
|---|---|---|---|---|
| | | Zn-peak | Zn-inflection | Zn-remote |
| Data collection | | | | |
| Wavelength(Å) | 1.2834 | 1.2834 | 1.2837 | 1.2155 |
| Space group | $P2_12_12_1$ | $P2_12_12_1$ | | |
| Resolution(Å) | 30.0–2.24 | 30.0–2.60 | 30.0–2.60 | 30.0–2.60 |
| Mean $I/\sigma(I)$ | 22.0 (5.8) | 23.2 (5.5) | 27.0 (3.7) | 27.5 (6.7) |
| Completeness (%) | 99.6 (97.8) | 94.2 (75.4) | 98.1 (97.6) | 96.6 (86.6) |
| <Redundancy> | 7.5 | 9.8 | 10.5 | 10.4 |
| Unique reflections | 15 928 | 11 884 | 12 008 | 12 403 |
| $R_{sym}$[a] (%) | 8.8 | 10.2 | 8.3 | 8.1 |
| Refinement statistics | | | | |
| Resolution (Å) | 20.00–2.24 | 15.00–2.60 | | |
| *R*-factor | 23.0 | 23.2 | | |
| $R_{free}$[b] | 26.4 | 28.2 | | |
| Average *B*-factors/no. of atoms | | | | |
| Protein (Å$^2$) | 41.9 | 62.9 | | |
| DNA (Å$^2$) | 46.4 | 61.4 | | |
| Zinc (Å$^2$)/zinc atoms | 45.4/4 | 58.6/4 | | |
| Solvent (Å$^2$)/solvent atoms | 54.6/233 | 58.8/102 | | |
| R.m.s deviations | | | | |
| Bond lengths (Å) | 0.008 | 0.011 | | |
| Bond angles (°) | 1.30 | 1.60 | | |
| Dihedral angles (°) | 20.5 | 21.6 | | |
| Improper angles (°) | 1.33 | 1.36 | | |

Numbers in parentheses are for the highest resolution shell.
[a]$R_{sym} = \Sigma|I_h - \langle I_h\rangle|/\Sigma I_h$, where $\langle I_h\rangle$ is the average intensity over symmetry-equivalent reflections.
[b]$R_{free}$ was calculated using 7% of the data excluded from refinement.

5′-AGGTCA-3′ consensus sequences, as shown in Figure 1D (Vögtli *et al.*, 1998; Niedziela-Majka *et al.*, 2000). The crystals of the EcR–USP-DBD heterodimer on DNA diffract to 2.24 Å resolution and contain a single complex in their asymmetric units.

The crystallization of the EcR–RXR/DNA complex relied on the same EcR protein construct (Figure 1A) and the same IR-1 DNA sequence (Figure 1E) used for the EcR–USP/IR-1 complex. The RXR-DBD used was an 80 residue DBD fragment derived from the human RXRα (residues 130–209), which contained five N-terminal residues preceding the N-terminal zinc-binding cysteine, the entire 66 residue core DBD and nine residues at the C-terminus beyond the GM amino acids (residues 65 and 66 in Figure 1C) that constitute the C-terminal boundary of the core DBD regions of all nuclear receptors. These crystals contain a single complex of EcR–RXR and IR-1 within their asymmetric units, and diffracted to 2.60 Å resolution. This structure was solved using multiple anomalous diffraction (MAD) phasing derived using the anomalous signal from the zinc ions of the DBDs. The refined coordinates were used subsequently as the search model to solve the structure of the EcR–USP/DNA complex (described above) by molecular replacement. For clarity, we rely on a common numbering scheme for EcR, USP and RXR from hereon, with residue numberings beginning with the N-terminal most cysteine coordinating the Zn-I module. Table I summarizes the crystallographic data and refinement statistics for both complexes.

Stereo views of both structures and the protein–DNA contacts are shown in Figure 2A and B, and a superposition of the EcR–USP/IR-1 and EcR–RXR/IR-1 complexes is shown in Figure 2C. Overall, the EcR–USP and the EcR–RXR complexes show an r.m.s.d. of 0.67 Å, when calculated over the entire DNA and all their Cα atoms. A side-by-side comparison of these two complexes indicates that the subunit arrangements, the protein–protein interactions that stabilize the heterodimer, and the protein–DNA interactions that stabilize each subunit on its respective half-site are generally well preserved in these two structures (Figures 2A and B, and 3A–C). Furthermore, the EcRE DNA structure in both complexes conforms to B-DNA geometry, lacking significant distortions, except at the spacer where there is similar minor groove narrowing in both structures (see Figure 4C).

Because the DNA is nearly symmetric, the same Zn-I and Zn-II modules of EcR and USP/RXR are brought to the subunit interface. These regions of EcR are distinct in sequence from their counterparts in USP/RXR, and thus the dimer interface uses different residues from each of the two DBDs despite the pseudo-symmetric interactions between the EcR and USP/RXR subunits. This observation is in contrast to those made in the case of GR–GR/IR-3 and ER–ER/IR-3 crystal structures, in which DBD interactions use identical residues from each subunit's Zn-II sites to form perfectly symmetric interactions (Luisi *et al.*, 1991; Schwabe *et al.*, 1993).

The A-box region in EcR, which is the C-terminal extension (CTE) of the core DBD, forms an ordered helical structure in the EcR–RXR complex, but this conformation is not seen in the EcR–USP complex where this region has disordered density instead. Similarly, the Zn-II regions in the EcR–USP/RXR complexes, which are observed here to contain short α-helices that foster dimerization and phosphate DNA binding, may assume different conformations in the uncomplexed

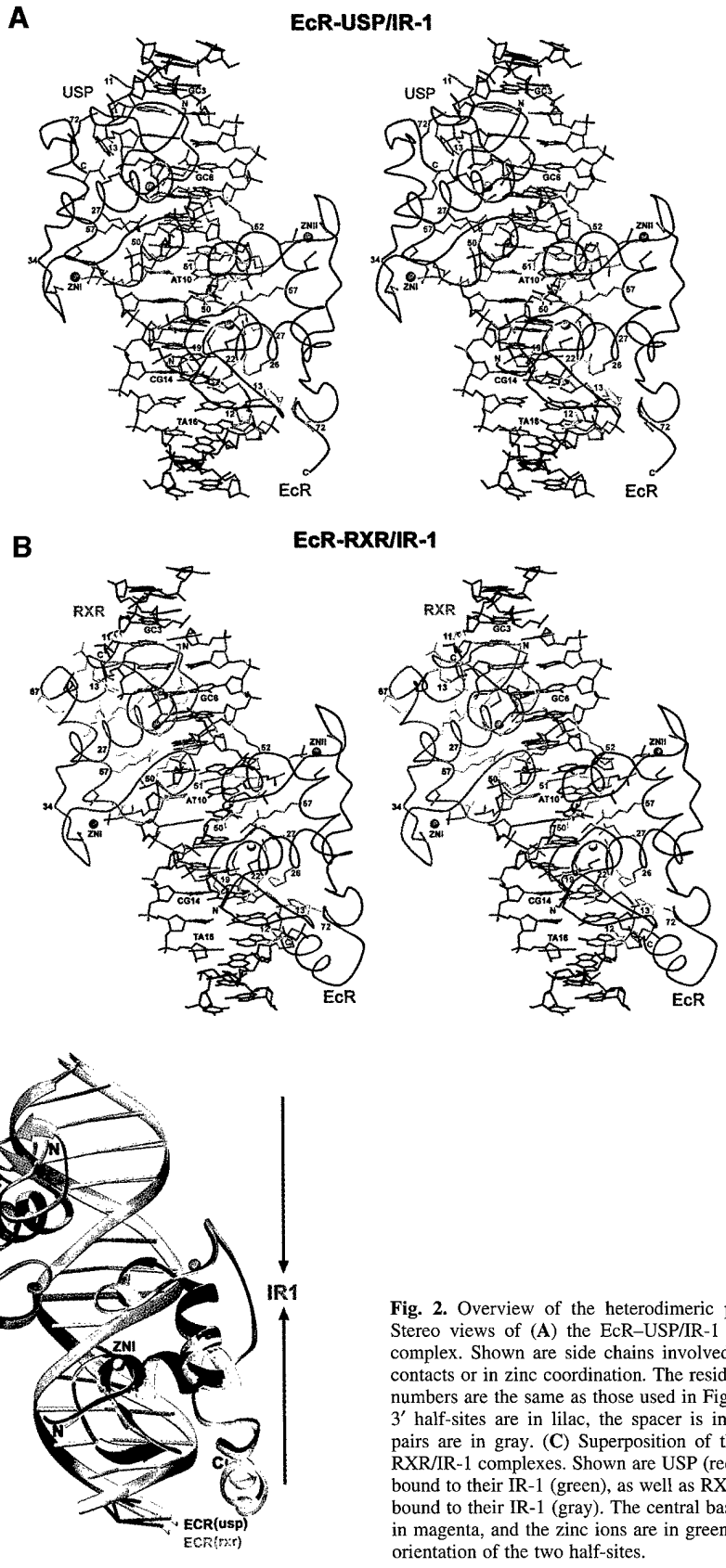



**Fig. 2.** Overview of the heterodimeric protein complexes on DNA. Stereo views of (**A**) the EcR–USP/IR-1 and (**B**) the EcR–RXR/IR-1 complex. Shown are side chains involved in direct base or phosphate contacts or in zinc coordination. The residue numbers and the base pair numbers are the same as those used in Figure 1. The two 5′-AGGTCA-3′ half-sites are in lilac, the spacer is in gold, and the flanking base pairs are in gray. (**C**) Superposition of the EcR–USP/IR-1 and EcR–RXR/IR-1 complexes. Shown are USP (red) and EcR (dark blue) DBDs bound to their IR-1 (green), as well as RXR (gold) and EcR (light blue) bound to their IR-1 (gray). The central base pair that acts as a spacer is in magenta, and the zinc ions are in green. Arrows indicate the relative orientation of the two half-sites.

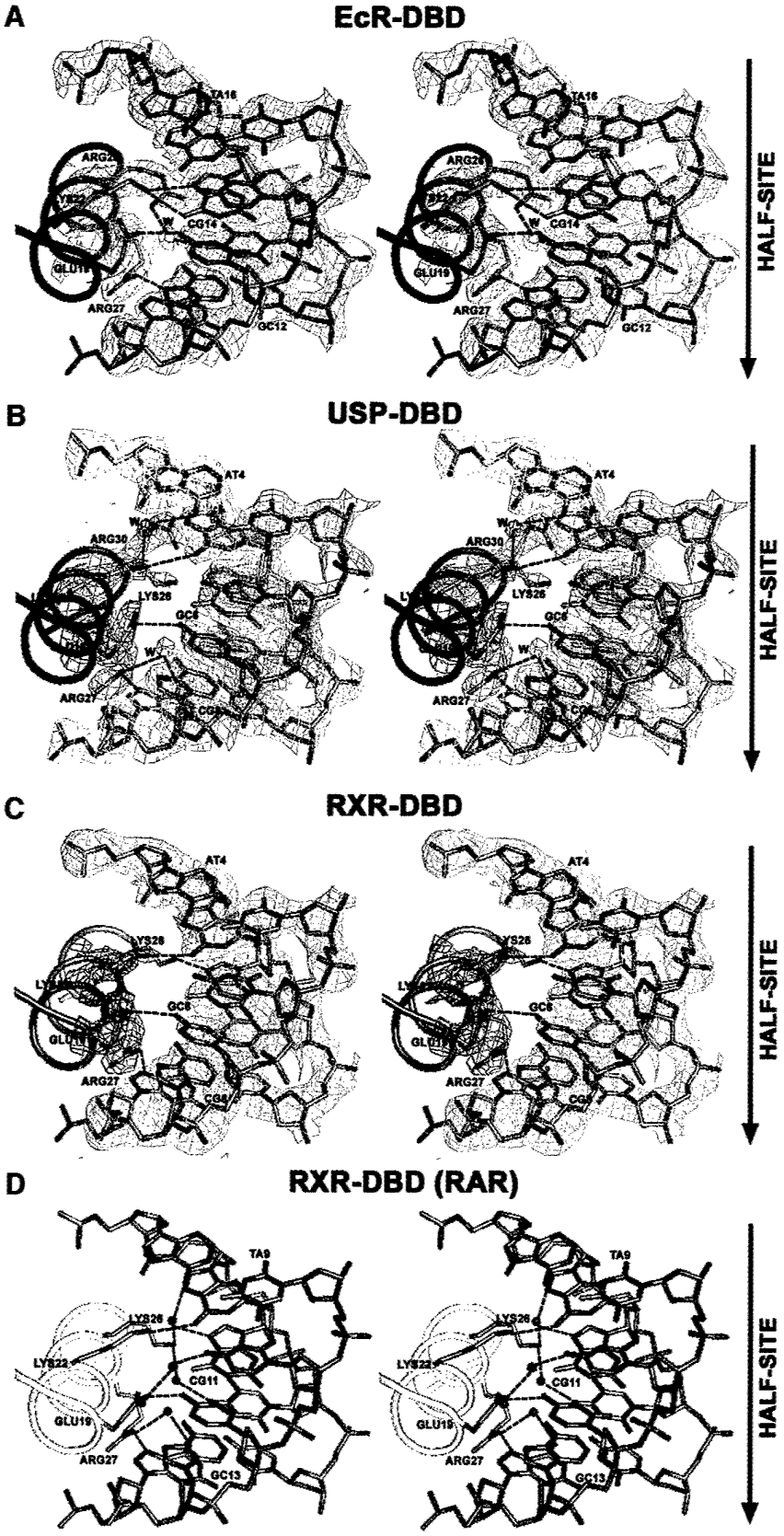



**Fig. 3.** Stereo views of the recognition α-helices and their contacts to DNA. Shown are interactions with the major grooves for (**A**) EcR (blue), (**B**) USP (red), (**C**) RXR (gold) with a composite omit electron density map ($2F_o - F_c$) shown for the side chains and the DNA base pairs, and (**D**) RXR from the RAR–RXR/DR-1 complex (yellow). In all the cases, the side chains that make direct or water-mediated contacts to the bases are shown in green, the water molecules that mediate the contacts between the protein and DNA are yellow/magenta spheres, and hydrogen bonds are shown as dotted lines. The orientation of the half-site is indicated by arrows.

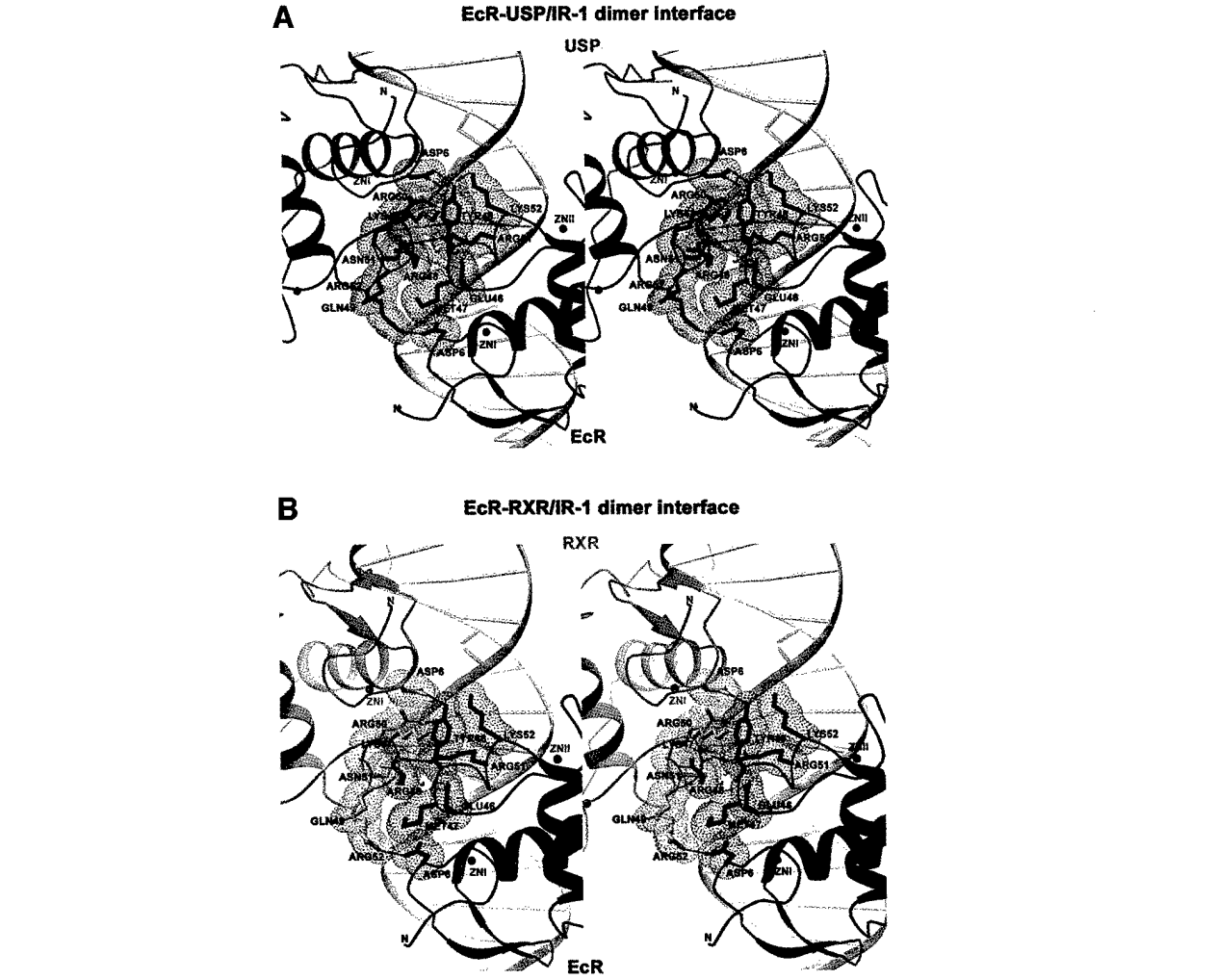
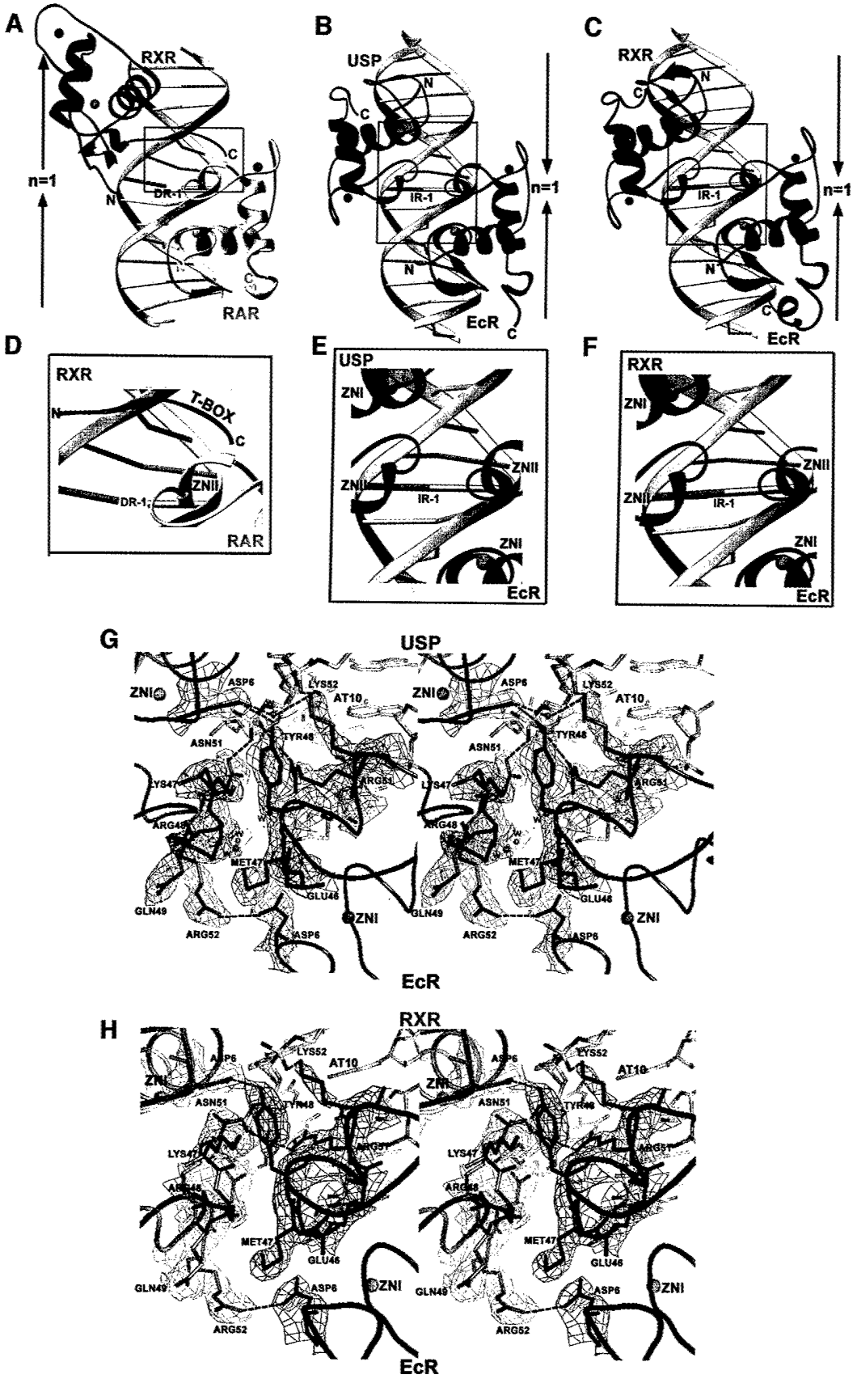



**Fig. 4.** The subunit interactions and role of DNA minor groove in dimerization. (**A**) Stereo view of the dimerization contacts between USP (red) and EcR (blue). (**B**) Stereo view of the dimerization contacts between RXR (gold) and EcR (blue). The side chains of the residues that make the dimer contacts are shown together with dotted surfaces indicating van der Waals complementary surfaces and dotted lines indicating hydrogen bonds. Also shown are the backbone DNA in green, the base pairs in gray and the spacer base pair in gold. The N- and C-termini of the two monomers are labeled, and the zincs are indicated by green spheres. (**C**) The minor groove width of the response element as a function of IR-1 sequence. The red and orange lines correspond to the EcR–USP and EcR–RXR DNAs, respectively. Both structures show a sharp minimum in the width at the spacer, the location in which interaction between the two proteins is supported. The minor groove width values were derived from the program CURVES (Lavery and Sklenar, 1988), which records the inter-phosphate distances after generating a smoothed curve through the phosphate atoms. (**D**) Stereo view of the electron density at the minor groove. Also shown are the contacts made by EcR to the phosphate backbone of the DNA together with the associated water molecules (W) located in the spine of hydration along the minor groove.


A
RXR
n=1
DR-1
RAR
B
USP
IR-1
n=1
EcR
C
RXR
IR-1
n=1
EcR
D
RXR
T-BOX
ZNII
DR-1
RAR
E
USP
ZNI
ZNII
IR-1
EcR
F
RXR
ZNI
ZNII
IR-1
EcR
G
USP
ZNI
ASP6
LYS52
AT10
ASN51
TYR48
LYS47
ARG48
MET47
GLU46
GLN49
ARG52
EcR
H
RXR
EcR

DBDs. This was the case observed in the GR, where similar helical conformations in the Zn-II regions were seen in the GR homodimer DBD complex on IR-3 DNA, but not seen in the monomeric GR-DBD off DNA (Luisi *et al.*, 1991; van Tilborg *et al.*, 1995).

### The protein–DNA contacts

In the EcR–USP complex, a total of 3110 Å$^2$ of water-accessible surface is buried in DNA binding, 1580 Å$^2$ of which is derived from the binding of the EcR subunit, and 1530 Å$^2$ from the binding of USP. In the EcR–RXR complex, a slightly larger area of 3250 Å$^2$ is buried, in part due to the contribution of the EcR C-terminal A-box helix, which is not ordered in the EcR–USP complex. These values are in line with those associated with DR-1 response element binding by RXR–RAR, which buries 3200 Å$^2$ (Rastinejad *et al.*, 2000), and involves similarly sized protein subunits, DNA sequences and 1 bp spacing (Figure 5A–C).

Interestingly, the pattern of hydrogen bonds between the recognition helices and the AGGTCA half-sites varies to some extent between EcR, USP, RXR and the previously characterized RXR subunit of the RXR–RAR/DR-1 heterodimer (Figure 3) (Rastinejad *et al.*, 2000). However, the same set of 'hook' residues are used consistently in all of these cases, with their exact hydrogen-bonded partners on the DNA base pairs differing slightly (Figure 3). These hook residues, consisting of Glu19, Gly20, Lys22, Arg/Lys26 and Arg27 in EcR, USP and RXR, form their base-specific contacts. Importantly, the central AT spacer base pair is involved in phosphate contacts only, and not major groove recognition at the base pairs (Figure 1D and E). This observation is consistent with the notion that the spacer's influence is limited to setting the correct binding site geometry in the response element (Ożyhar and Pongs, 1993).

### The dimer interfaces

Dimerization of DBDs is of major functional significance for RXR, its partners and the steroid receptors, in part because dimerization allows the formation of sufficiently extended and complementary protein surfaces for high affinity and selective DNA binding. Figures 1, 4, and 5G and H show how binding of one subunit at its half-site leads to facilitated binding of the adjacent subunit via protein–protein interactions. Due to the single base pair spacer, the relative orientation between the subunits is distinct from complexes involving steroid receptors GR and ER, and as such brings into contact the Zn-I modules of both proteins in addition to their Zn-II modules. Moreover, there are clear distinctions with previously studied heterodimers of RXR-DBD on direct repeats, in which the upstream subunit relies on its Zn-II region and the downstream subunit relies on its T- and/or A-boxes for protein–protein contacts (Figure 5A–H) (Khorasanizadeh and Rastinejad, 2001; Rastinejad, 2001).

In the current structures, neither subunit relies on its T- or A-boxes for dimerization, with these regions instead being pointed towards the 3′- and 5′-flanking sequences of the IR-1 element. The involvement of USP T-box sequence in interactions with 5′-flanking sequence has been suggested previously by Grad *et al.* (2001) who analyzed interactions of derivatives of the natural 20E response element with a USP-DBD protein in which the C-terminal region was deleted. However, a deletion removing the A-box of EcR disrupted the DNA binding of the EcR–USP (Niedziela-Majka *et al.*, 2000), which cannot be explained by our crystallographic results. Any additional role for the EcR A-box in dimerization cannot be ruled out, as our structure allows us to visualize only seven out of 25 A-box residues.

In the EcR–USP complex, a total of 560 Å$^2$ of water-accessible surface area is buried between the subunits, whereas all previously studied DBD homo- and heterodimers bury a smaller surface, in the range 300–480 Å$^2$. Figures 4A and B, and 5G and H show the detailed dimer interface in the EcR–USP and EcR–RXR complexes, respectively. These dimerization interfaces rely on hydrogen bonds between subunit residue side chains as well as significant complementary surfaces relying on van der Waals contacts. Essentially all the contacts between the subunits are conserved between the EcR–USP and EcR–RXR complexes.

### Cooperation between EcR and USP/RXR

Nuclear receptor DBDs do not form homo- or heterodimers in the absence of DNA (Hard *et al.*, 1990; Mader *et al.*, 1993; Perlmann *et al.*, 1993; Schwabe *et al.*, 1993; Zechel *et al.*, 1994a). Receptor homo- or heterodimer formation through DBDs is strictly dependent and enhanced by the cognate DNA-binding sites. In the heterodimeric complexes studied here, the structures suggest that there are three mechanisms by which the IR-1 appears cooperatively to enhance the dimerization between the EcR and the USP/RXR homologs. First, the same Zn-II regions involved in the formation of the dimer interface are also used extensively for making DNA contacts (Figures 1A–C, and 5G and H). In particular, residues Arg51 and Lys52 from EcR and residue Asn51 of USP are simultaneously involved in both dimerization and DNA binding functions. This implies that DNA binding and subunit dimerization are mutually supportive.

A second mechanism by which the DNA enhances the dimer interactions can be seen in Figures 4A and B, and 5G and H, showing how the subunit interfaces are in part embedded in the minor groove. Figure 4C and D shows that a significant minor groove distortion is associated with the spacer AT base pair, this being the convergence point of the protein–protein interactions. Importantly, these minor groove widths represent sharp departures from standard B-DNA values, and are associated with both the EcR–USP and EcR–RXR structures. In particular, there is

**Fig. 5.** Comparison of the RXR–RAR complex on DR-1 with those of EcR–USP and EcR–RXR on IR-1. Side-by-side comparison of (**A**) RAR–RXR/DR-1 with (**B**) EcR–USP/IR-1 and (**C**) EcR–RXR/IR-1 complexes. The arrows indicate the relative orientation of the half-sites associated with inverted repeat and direct repeat response elements. Boxed regions are shown in close-up views to indicate the dimer interfaces in (**D**) RAR–RXR/DR-1, (**E**) EcR–USP/IR-1 and (**F**) EcR–RXR/IR-1. Also shown are stereo views of the composite omit ($2F_o - F_c$) electron density map at the dimer interface of (**G**) EcR–USP/IR-1 and (**H**) EcR–RXR/IR-1.

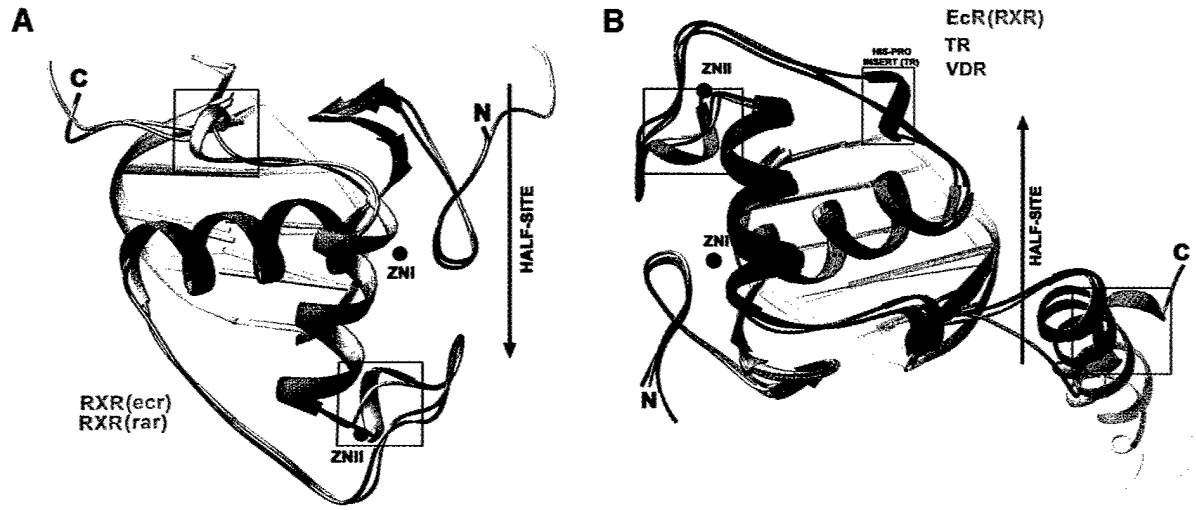



**Fig. 6.** Comparison of EcR complexes with other DBD–DNA complexes. (**A**) RXR on its 5′-AGGTCA-3′ site from EcR–RXR/IR-1 is superimposed onto RXR on its downstream 5′-AGGTCA-3′ half-site from RXR–RAR/DR-1. (**B**) EcR from the EcR–RXR/IR-1 complex is superimposed onto TR (TR–RXR/DR-4) and VDR (VDR–VDR/DR-3), the latter of which are from downstream half-sites in direct-repeats. While the overall fold of the receptors and their placement of the half-site DNA elements are very similar, the notable differences (indicated by boxes) can be attributed to the use of some protein regions for dimerization contacts.

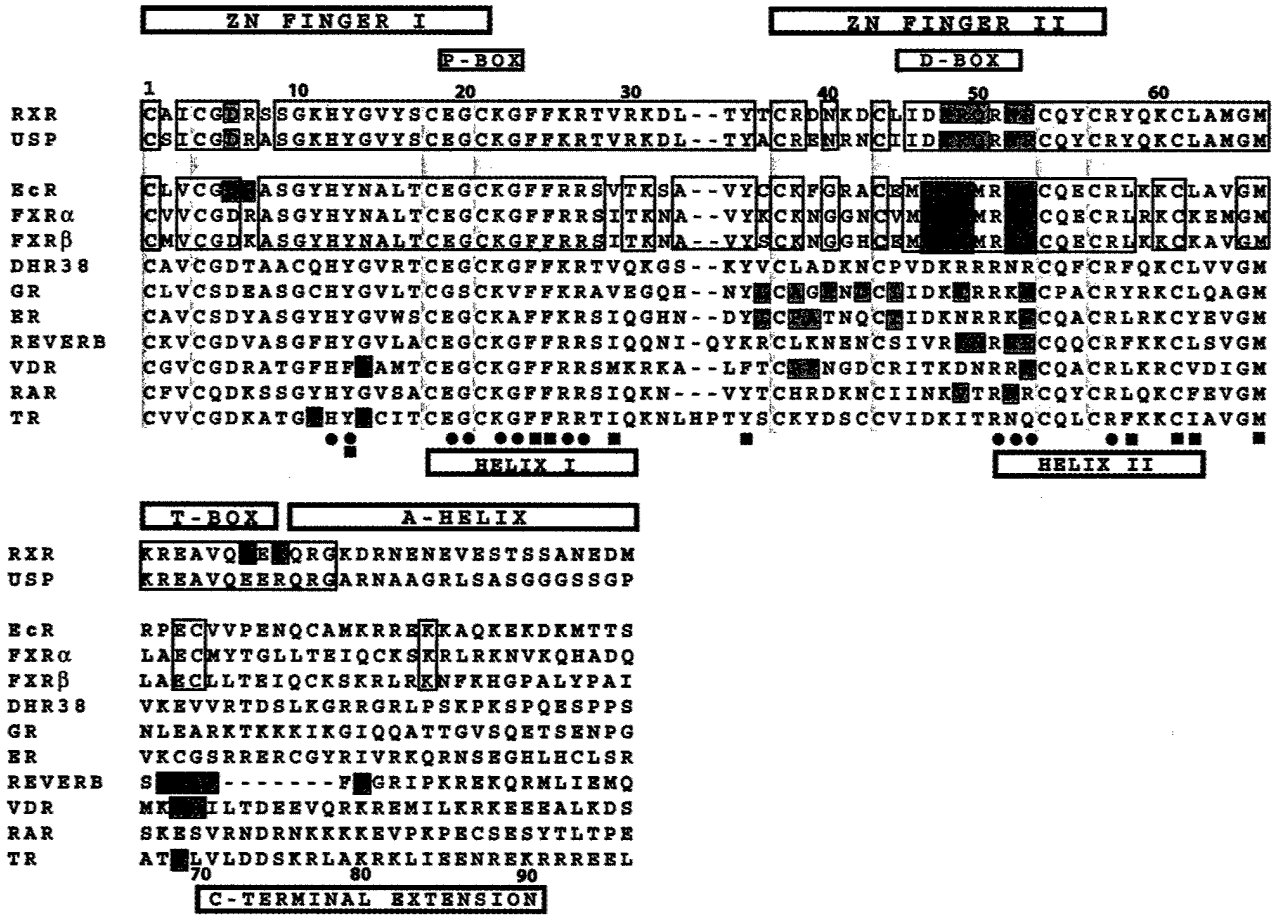



**Fig. 7.** Sequence comparison of nuclear receptor DBDs and the conservation of their protein–protein and protein–DNA contacts. The zinc-coordinating cysteines are in yellow, and the regions corresponding to the Zn-I, Zn-II, helix I (recognition helix), helix-II and CTEs are labeled. Blue dots indicate residues that are involved in specific DNA binding, and red squares are the residues that form the hydrophobic core that stabilizes the protein. The residues involved in dimer contacts are shaded in blue (RXR and USP) and green (EcR). Importantly, all of the DNA-binding residues are fully conserved between RXR and USP, and between EcR, and FXRα and FXRβ.

a 4.3 Å minor groove width in the EcR–USP DNA and a <4.0 Å width in the EcR–RXR DNA. The reliance on minor groove distortions to stabilize dimer binding is reminiscent of the RXR–RAR/DR-1 and the RevErb–RevErb/DR-2 complexes on their cognate DNA targets (Rastinejad *et al.*, 2000; Sierk *et al.*, 2001).

A third mechanism for subunit cooperation is evident in Figure 1D and E, which shows that the EcR-DBD footprint on DNA extends well beyond its own AGGTCA site to reach over both its 3′-flanking sequences and a large portion of the USP half-site. In total, the EcR footprint in the USP complex extends over a region totaling 13 bp, and to 12 bp in the RXR complex (Figure 1D and E). This is consistent with mutational studies which identified base pairs within the USP half-site that were found to be critical not for the USP-DBD binding (as a monomer) but for effective heterodimer formation (Grad *et al.*, 2001). This extended binding mode exhibited by EcR may contribute to the cooperativity of subunit association, by reducing the conformational flexibility at the USP site and as such pre-paying the entropic costs associated with adjacent site binding by USP. A similar mechanism based on tandem site stabilization has been suggested as the basis for the cooperation between the POU domains of Oct-1 on DNA, as well as the binding of the RXR–RAR heterodimer on DNA (Klemm and Pabo, 1996; Rastinejad *et al.*, 2000).

### *Comparison with other nuclear receptor DBDs*

Figure 5A–F shows how RXR-DBD and USP-DBD, derived from the RXR–RAR, EcR–USP and EcR–RXR crystal structures, form their respective dimerization contacts on DR-1 and IR-1 elements. Importantly, the C-terminal extension (T-box) of the RXR-DBD, used extensively in heterodimeric interactions with RAR on the DR-1 complex, is not used similarly in the EcR complex on IR-1 (Figure 5D–F). Moreover, the Zn-I and Zn-II regions used in both the EcR–USP and EcR–RXR complexes on IR-1 are not used for heterodimerization with RAR on DR-1. Therefore, RXR/USP are able to use distinct regions of their DBDs for forming dimerization contacts on DR-1 and IR-1. These crystallographic observations are supported by previously reported biochemical studies that rely on gel shift experiments with RXR, RAR, USP and EcR DBDs (Lee *et al.*, 1993; Zechel *et al.*, 1994a, b; Niedziela-Majka *et al*, 2000). Specifically, when the T-box residues of USP (KREAVQEERQR) are deleted from the DBD, USP is still able to form its cooperative interactions with EcR on IR-1 (Niedziela-Majka *et al.*, 2000), as predicted from the structures shown in Figure 5B, C, E and F. However, a deletion within the T-box of mRXR (removing the residues EERQR and beyond) prevents RXR from forming a protein–protein interaction with RAR on DR-1 (Zechel *et al.*, 1994a, b), consistent with its importance in the RXR–RAR/DR-1 complex (Figure 5A and D). A similar T-box mutation in hRXR-DBD (removing the sequences VQEERQR and beyond) also diminishes the ability of RXR to form cooperative homodimers on DR-1 (Lee *et al.*, 1993).

To see how the RXR-DBD subunits, derived from the EcR–RXR and RAR–RXR DBD complexes (Figure 5A–C) adjust their conformations to adapt their partners and DNA elements, we superimposed their backbone structures on half-sites in Figure 6A. The boxed areas in Figure 6A show the major differences in conformation, which correspond to two regions both of which are involved in the formation of their respective heterodimeric complexes. These are the C-terminal extension of the DBD, used for heterodimeric interactions with RAR on the DR-1 complex, and the Zn-II region used in heterodimeric interactions in the EcR complex. Interestingly, the T-box of RXR, which has also been shown to be an α-helix in the absence of DNA (Holmbeck *et al.*, 1998), adopts a conformation not seen in the current structures or in any of the RXR–RXR and RXR–RAR complexes previously observed bound to DR-1. These differences provide further evidence that adaptable surfaces in nuclear receptor DBDs readily undergo structural rearrangements that help accommodate their association with their dimerization partners and response elements.

As noted above, the EcR subunit associated with the RXR complex described here showed an α-helical structure in its A-box region. To see how the EcR-DBD compares with other steroid and non-steroid DBD structures previously reported to contain α-helical A-boxes within their C-terminal extensions, we superimposed their half-complexes in Figure 6B. This comparison used only the VDR (from the VDR homodimer) (Shaffer and Gewirth, 2002) and TR (from TR–RXR) (Rastinejad *et al.*, 1995), as similar helical regions had not been observed in other DBD co-crystal structures. The VDR and TR subunits, which are both positioned in the downstream half-site of direct repeat complexes, use their respective C-terminal helices to form protein–protein and protein–DNA interfaces that stabilize their respective complexes. In the case of EcR, this helical region is not near the dimer interface and is instead directed 5′ to the half-site. Moreover, the helical region was visualized here only in the EcR–RXR complex and not in the EcR–USP complex.

## Discussion

While the binding of a heterodimer to a symmetric response element has been observed with other transcription factors, this form of subunit association has never before been observed in the nuclear receptor family. As there is a high level of amino acid sequence conservation between RXR and USP DBDs, together with the observation that RXR and USP can substitute for each other in DNA binding with EcR (Yao *et al.*, 1992; Vögtli *et al.*, 1998), we also studied the crystal structure of the EcR–RXR/IR-1 complex. USP and RXR differ in only six residues, and our findings indicate that structural determinants for EcR and DNA binding are well conserved in USP and RXR (Figure 7).

Like EcR, FXR also forms heterodimers with RXR that bind to IR-1 sites (Forman *et al.*, 1995; Laffitte *et al.*, 2000). Our phylogenetic analysis, using members of the superfamily and focused strictly on the portion of receptors including the DBDs and 30 residues at the CTE, indicate that EcR is most closely related to FXR. Within the core DBD of 66 residues, which contains all of the determinants of EcR binding to DNA and USP, there are only eight non-conserved residues between EcR and FXR (Figure 7). None of the non-conserved residues in FXR fall at sites used for response element binding or

subunit dimerization by EcR in its IR-1 complexes. Moreover, Figure 7 shows that the critical interacting residues in EcR are equally shared with both hFXRα and mFXRβ, but not with other receptors. Therefore, it is likely that the current structure of EcR–RXR closely mimics the mode of subunit interactions used by the FXR–RXR/IR-1 complex.

In addition, the *Drosophila* receptor DHR38 is also known to form heterodimers with USP (Henrich *et al.*, 1994; Sutherland *et al.*, 1995; Baker *et al.*, 2003). However, DHR38 appears much more closely related to the human NGFI-B than to either EcR or FXR. The sequence alignment in Figure 7 shows that the determinants of dimerization with USP used by EcR are not at all conserved in DHR38. These findings would suggest that USP–DHR38 either use other types of DNA response elements, or form an altogether different set of protein–protein interactions than described here in the EcR complexes. Figure 7 also shows that while the known determinants of DBD dimerization in GR, ER, RevErb, VDR, RAR and TR involve to some extent the Zn-II regions, the residues used differ in sequence and position.

The nuclear receptor USP is required in multiple tissues of the organism at various stages of metamorphosis. All these factors together implicate USP as a regulator of multiple pathways in *Drosophila*. It is also known that mutations in USP result in complex embryonic and adult phenotypes, making it a vital factor involved in *Drosophila* development. This pleiotropy could be based on the interaction of USP with other factors involved in development, reminiscent of the role of mammalian RXR in regulating multiple pathways. It is interesting to note that while these two receptor homologs do not share the same ligand-binding properties, they nevertheless each act as a common dimerization partner for multiple other nuclear receptors in flies or mammals.

Because EcR relies on both an inverted repeat target and heterodimerization, this receptor has characteristics typical of both steroid and non-steroid receptors, and may be a clear evolutionary link between vertebrate steroid and non-steroid receptors. Apart from the degenerate palindromic elements, the EcR dimer also activates transcription *in vivo* through direct repeats (D'Avino *et al.*, 1995). Since it is evident that EcR is functional only in the presence of USP, this implies that the two subunits are capable of dimerizing on a variety of response elements. At a physiological level, this versatility could be the mechanism through which they regulate multiple pathways. At a molecular level, this further supports the notion that flexible and adjustable dimerization surfaces in DBDs are responsible for their cooperation on various response elements.

## Materials and methods

### *Crystallization and data collection*

The DBDs of EcR, USP and RXR were expressed in *Escherichia coli* as fusions with GST, using pGEX-2T (USP, EcR) and pGEX-4T (RXR) vectors (Pharmacia) (Rastinejad *et al.*, 1995; Niedziela-Majka *et al.*, 1998). The proteins and the two oligonucleotide strands corresponding to IR-1 were purified as described previously (Zhao *et al.*, 2000; Niedziela-Majka *et al.*, 2001). Samples of EcR–RXR co-crystallization contained DNA and the two proteins at concentrations of 0.5 mM each. Crystals were grown using hanging drops at 9°C by addition of 2 µl of the complex to an equal volume of a solution containing 12–14% PEG 3350, 50 mM $MgSO_4$, 50 mM $Li_2SO_4$, 5 mM dithiothreitol (DTT), 10 mM $MgCl_2$, 0.1 M MES pH 5.6. Crystals were streaked through a cryo-solvent solution containing the reservoir solution supplemented with 20% glycerol, and flash frozen in liquid nitrogen. Diffraction data were collected on these frozen crystals under a stream of liquid nitrogen at 100 K, at the Argonne National Laboratory at beamline SBC-CAT 19ID. The DBDs contain two zinc ions each, which were used to obtain MAD data at three different wavelengths (1.2834, 1.2837 and 1.2155 Å) corresponding to the zinc peak, inflection and a remote wavelength. The crystals diffracted to a resolution of 2.60 Å and crystallized in the orthorhombic space group ($P2_12_12_1$, $a = 55.62$, $b = 60.25$, $c = 115.00$).

The EcR–USP complex was crystallized with the same IR-1 response element. Crystals were grown using hanging drops at 9°C by addition of 2 µl of the complex (0.5 mM) to an equal volume of a solution containing 4–8% PEG 3350, 100 mM NaCl, 5 mM DTT, 10 mM $MgCl_2$ and 0.1 M MES pH 5.6. Crystals were streaked through a cryo-solvent solution containing the reservoir solution containing 20% glycerol, and flash frozen in liquid nitrogen. X-ray diffraction data were collected on these frozen crystals under a stream of liquid nitrogen, at the BNL-X4A beamline. The crystals diffracted to 2.24 Å and crystallized in orthorhombic space group ($P2_12_12_1$, $a = 50.37$, $b = 59.98$, $c = 113.57$). All the data collected were integrated and scaled using HKL2000 (Otwinowski and Minor, 1997).

### *Structure determination*

Phases were calculated from the data obtained from the EcR–RXR crystals using SOLVE (Terwilliger and Berendzen, 1999) and resulted in a figure of merit of 0.61 (for data to 3.0 Å resolution). SOLVE identified the positions of all four zinc ions in the complex. RESOLVE and DM (CCP4) were used for solvent flattening and density modification, improving the quality of the map. These phases were used to calculate an initial 3.0 Å electron density map, which was readily interpretable in terms of the DNA duplex and the backbone chains of both the proteins. A partial model consisting of the core DBDs of both the proteins (with side chains for the conserved residues) and most of the DNA duplex was built using the program O (Jones *et al.*, 1991). The model was partially refined in CNS (Brünger *et al.*, 1998) by rigid body refinement, simulated annealing and energy minimization. The refined model provided improved maps which made it possible to distinguish between the two proteins. All the side chains in the core DBDs were gradually added and the model was refined to *R*-values of 38 and 32% ($R_{free}$ and *R*-factor). This model was used as the search model for molecular replacement to obtain the EcR–USP structure. The entire complex including the zinc atoms, the two subunits and the DNA was used as the search model for molecular replacement. Molecular replacement was carried out using MOLREP (CCP4) and gave a clear solution with a correlation coefficient of 0.40 and an *R*-factor of 45%.

Both models were refined in CNS with successive rounds of rigid-body refinement, simulated annealing and energy minimization (Brünger *et al.*, 1998). Initially tight restraints were imposed for Watson–Crick DNA base pairing and the planarity of the atoms in the bases (Parkinson *et al.*, 1996). In addition, all the residues in the 5′-subunit in the EcR–RXR/IR-1 model were changed to correspond to those in USP. Improved maps helped visualize residues beyond the core DBD in all the proteins. The tight restraints subsequently were released, followed by multiple rounds of manual rebuilding and refinement in CNS including individual *B*-factor refinement. Solvent molecules were added to the model using Arp waters (Cowtan, 1994), and the model refined to *R*-values as shown in Table I.

## Acknowledgements

We thank Christine Wright and Li-Zhi Mi for helpful discussions, Anita Niedziela-Majka and Iwona Grad for sharing reagents, and Craig Ogata and Randy Abramowitz for assistance with data collection at BNL-X4A. This study was supported by NIH grant GM55217 (F.R.), a State Committee for Scientific Research grant (Poland) 3 P04B 009 23 (A.O.), and a fellowship from the Rett Syndrome Research Foundation (J.M.H.).

## References

Antoniewski,C., Laval,M. and Lepesant,J.A. (1993) Structural features critical to the activity of an ecdysone receptor binding site. *Insect Biochem. Mol. Biol.*, **23**, 105–114.

Antoniewski,C., Laval,M., Dahan,A. and Lepesant,J.A. (1994) The

ecdysone response enhancer of the Fbp1 gene of *Drosophila melanogaster* is a direct target for the EcR/USP nuclear receptor. *Mol. Cell. Biol.*, **14**, 4465–4474.

Antoniewski,C., Mugat,B., Delbac,F. and Lepesant,J.A. (1996) Direct repeats bind the EcR/USP receptor and mediate ecdysteroid responses in *Drosophila melanogaster*. *Mol. Cell. Biol.*, **16**, 2977–2986.

Ashburner,M. (1973) Sequential gene activation by ecdysone in polytene chromosomes of *Drosophila melanogaster*. I. Dependence upon ecdysone concentration. *Dev. Biol.*, **35**, 47–61.

Ashburner,M. (1974) Sequential gene activation by ecdysone in polytene chromosomes of *Drosophila melanogaster*. II. The effects of inhibitors of protein synthesis. *Dev. Biol.*, **39**, 141–157.

Baker,K.D. *et al.* (2003) The *Drosophila* orphan nuclear receptor DHR38 mediates an atypical ecdysteroid signaling pathway. *Cell*, **113**, 731–742.

Brünger,A.T. *et al.* (1998) Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D*, **54**, 905–921.

Chawla,A., Repa,J.J., Evans,R.M. and Mangelsdorf,D.J. (2001) Nuclear receptors and lipid physiology: opening the X-files. *Science*, **294**, 1866–1870.

Cherbas,L., Lee,K. and Cherbas,P. (1991) Identification of ecdysone response elements by analysis of the *Drosophila* Eip28/29 gene. *Genes Dev.*, **5**, 120–131.

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

D'Avino,P.P., Crispi,S., Cherbas,L., Cherbas,P. and Furia,M. (1995) The moulting hormone ecdysone is able to recognize target elements composed of direct repeats. *Mol. Cell. Endocrinol.*, **113**, 1–9.

Fisk,G.J. and Thummel,C.S. (1995) Isolation, regulation and DNA-binding properties of three *Drosophila* nuclear hormone receptor superfamily members. *Proc. Natl Acad. Sci. USA*, **92**, 10604–10608.

Forman,B.M. *et al.* (1995) Unique response pathways are established by allosteric interactions among nuclear hormone receptors. *Cell*, **81**, 687–693.

Gilbert,L.I., Song,Q. and Rybczynski,R. (1997) Control of ecdysteroidogenesis: activation and inhibition of prothoracic gland activity. *Invert. Neurosci.*, **3**, 205–216.

Gilbert,L.I., Rybczynski,R. and Warren,J.T. (2002) Control and biochemical nature of the ecdysteroidogenic pathway. *Annu. Rev. Entomol.*, **47**, 883–916.

Glass,C.K. (1994) Differential recognition of target genes by nuclear receptor monomers, dimers and heterodimers. *Endocr. Rev.*, **15**, 391–407.

Grad,I., Niedziela-Majka,A., Kochman,M. and Ożyhar,A. (2001) Analysis of Usp DNA binding domain targeting reveals critical determinants of the ecdysone receptor complex interaction with the response element. *Eur. J. Biochem.*, **268**, 3751–3758.

Hall,B.L. and Thummel,K.E. (1998) The RXR homolog ultraspiracle is an essential component of the *Drosophila* ecdysone receptor. *Development*, **125**, 4709–4717.

Hard,T. *et al.* (1990) Solution structure of the glucocorticoid receptor DNA-binding domain. *Science*, **249**, 157–160.

Henrich,V.C., Szekely,A.A., Kim,S.J., Brown,N.E., Antoniewski,C., Hayden,M.A., Lepesant,J.A. and Gilbert,L.I. (1994) Expression and function of the ultraspiracle (usp) gene during development of *Drosophila melanogaster*. *Dev. Biol.*, **165**, 38–52.

Holmbeck,S.M., Dyson,H.J. and Wright,P.E. (1998) High-resolution solution structure of the retinoid X receptor DNA-binding domain. *J. Mol. Biol.*, **284**, 533–539.

Horner,M.A., Chen,T. and Thummel,C.S. (1995) Ecdysteroid regulation and DNA binding properties of *Drosophila* nuclear hormone receptor superfamily members. *Dev. Biol.*, **168**, 490–502.

Jones,T.A., Zou,J.Y., Cowan,S.W. and Kjeldgaard,M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A*, **47**, 110–119.

Khorasanizadeh,S. and Rastinejad,F. (2001) Nuclear-receptor interactions on DNA-response elements. *Trends Biochem. Sci.*, **26**, 384–390.

Klemm,J.D. and Pabo,C.O. (1996) Oct-1 POU domain–DNA interactions: cooperative binding of isolated subdomains and effects of covalent linkage. *Genes Dev.*, **10**, 27–36.

Koelle,M.R., Talbot,W.S., Segraves,W.A., Bender,M.T., Cherbas,P. and Hogness,D.S. (1991) The *Drosophila EcR* gene encodes an ecdysone receptor, a new member of the steroid receptor superfamily. *Cell*, **67**, 59–77.

Laffitte,B.A., Kast,H.R., Nguyen,C.M., Zavacki,A.M., Moore,D.D. and Edwards,P.A. (2000) Identification of the DNA binding specificity and potential target genes for the farnesoid X-activated receptor. *J. Biol. Chem.*, **275**, 10638–10647.

Lavery,R. and Sklenar,H. (1988) The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids. *J. Biomol. Struct. Dyn.*, **6**, 63–91.

Lee,M.S., Kliewer,S.A., Provencal,J., Wright,P.E. and Evans,R.M. (1993) Structure of the retinoid X receptor α DNA binding domain: a helix required for homodimeric DNA binding. *Science*, **260**, 1117–1121.

Lehmann,M. and Korge,G. (1995) Ecdysone regulation of the *Drosophila* Sgs-4 gene is mediated by the synergistic action of ecdysone receptor and SEBP 3. *EMBO J.*, **14**, 716–726.

Lehmann,M., Wattler,F. and Korge,G. (1997) Two new regulatory elements controlling the *Drosophila* Sgs-3 gene are potential ecdysone receptor and fork head binding sites. *Mech. Dev.*, **62**, 15–27.

Luisi,B.F., Xu,W.X., Otwinowski,Z., Freedman,L.P., Yamamoto,K.R. and Sigler,P.B. (1991) Crystallographic analysis of the interaction of the glucocorticoid receptor with DNA. *Nature*, **352**, 497–505.

Mader,S., Chen,J.Y., Chen,Z., White,J., Chambon,P. and Gronemeyer,H. (1993) The patterns of binding of RAR, RXR and TR homo- and heterodimers to direct repeats are dictated by the binding specificites of the DNA binding domains. *EMBO J.*, **12**, 5029–5041.

Mangelsdorf,D.J. and Evans,R.M. (1995) The RXR heterodimers and orphan receptors. *Cell*, **83**, 841–850.

Mangelsdorf,D.J. *et al.* (1995) The nuclear receptor superfamily: the second decade. *Cell*, **83**, 835–839.

Mi,L.Z., Devarakonda,S., Harp,J.M., Han,Q., Pellicciari,R., Willson,T.M., Khorasanizadeh,S. and Rastinejad,F. (2003) Structural basis for bile acid binding and activation of the nuclear receptor FXR. *Mol. Cell*, **11**, 1093–1100.

Niedziela-Majka,A., Rymarczyk,G., Kochman,M. and Ożyhar,A. (1998) Pure, bacterially expressed DNA-binding domains of the funtional ecdysteroid receptor capable of interacting synergestically with the hsp 27 20-hydroxyecdysone response element. *Protein Expr. Purif.*, **14**, 208–220.

Niedziela-Majka,A., Kochman,M. and Ożyhar,A. (2000) Polarity of the ecdysone receptor complex interaction with the palindromic response element from the *hsp27* gene promoter. *Eur. J. Biochem.*, **267**, 507–519.

No,D., Yao,T.P. and Evans,R.M. (1996) Ecdysone-inducible gene expression in mammalian cells and transgenic mice. *Proc. Natl Acad. Sci. USA*, **93**, 3346–3351.

Oro,A.E., McKeown,M. and Evans,R.M. (1990) Relationship between the product of the *Drosophila* ultraspiracle locus and the vertebrate retinoid X receptor. *Nature*, **347**, 298–301.

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

Ożyhar,A. and Pongs,O. (1993) Mutational analysis of the interaction between ecdysteroid receptor and its response element. *J. Steroid Biochem. Mol. Biol.*, **46**, 135–145.

Ożyhar,A., Strangmann-Diekmann,M., Kiltz,H.H. and Pongs,O. (1991) Characterization of a specific ecdysteroid receptor–DNA complex reveals common properties for invertebrate and vertebrate hormone-receptor/DNA interactions. *Eur. J. Biochem.*, **200**, 329–335.

Parkinson,G., Wilson,C., Gunasekera,A., Ebright,Y.W., Ebright,R.E. and Berman,H.M. (1996) Structure of the CAP–DNA complex at 2.5 angstroms resolution: a complete picture of the protein–DNA interface: atomic hydrogen bond in sequence-specific protein DNA recognition. *J. Mol. Biol.*, **260**, 395–408.

Perlmann,T., Rangarajan,P.N., Umesono,K. and Evans,R.M. (1993) Determinants for selective RAR and TR recognition of direct repeat HREs. *Genes Dev.*, **7**, 1411–1422.

Rastinejad,F. (2001) Retinoid X receptor and its partners in the nuclear receptor family. *Curr. Opin. Struct. Biol.*, **11**, 33–38.

Rastinejad,F., Perlmann,T., Evans,R.M. and Sigler,P.B. (1995) Structural determinants of nuclear receptor assembly on DNA direct repeats. *Nature*, **375**, 203–211.

Rastinejad,F., Wagner,T., Zhao,Q. and Khorasanizadeh,S. (2000) Structure of the RXR–RAR DNA-binding complex on the retinoic acid response element DR1. *EMBO J.*, **19**, 1045–1054.

Riddiford,L.M. (1993) Hormone receptors and the regulation of insect metamorphosis. *Receptor*, **3**, 203–209.

Riddiford,L.M. (1996) Juvenile hormone: the status of its 'status quo' action. *Arch. Insect Biochem. Physiol.*, **32**, 271–286.

Riddihough,G. and Pelham,H.R.B. (1987) An ecdysone response element in the *Drosophila* hsp27 promoter. *EMBO J.*, **6**, 3729–3734.

Schwabe,J.W., Chapman,L., Finch,J.T. and Rhodes,D. (1993) The

crystal structure of the estrogen receptor DNA-binding domain bound to DNA: how receptors discriminate between their response elements. *Cell*, **75**, 567–578.

Shaffer,P.L. and Gewirth,D.T. (2002) Structural basis of VDR–DNA interactions on direct repeat response elements. *EMBO J.*, **21**, 2242–2252.

Sierk,M.L., Zhao,Q. and Rastinejad,F. (2001) DNA deformability as a recognition feature in the reverb response element. *Biochemistry*, **40**, 12833–12843.

Sutherland,J.D., Kozlova,T., Tzertzinis,G. and Kafatos,F.C. (1995) *Drosophila* hormone receptor 38: a second partner for *Drosophila* USP suggests an unexpected role for nuclear receptors of the nerve growth factor-induced protein B type. *Proc. Natl Acad. Sci. USA*, **92**, 7966–7970.

Terwilliger,T.C. and Berendzen,J. (1999) Automated MAD and MIR structure solution. *Acta Crystallogr. D*, **55**, 849–861.

Thomas,H.E., Stunnenberg,H.G. and Stewart,A.F. (1993) Heterodimerization of the *Drosophila* ecdysone receptor with retinoid X receptor and ultraspiracle. *Nature*, **362**, 471–475.

Umesono,K. and Evans,R.M. (1989) Determinants of target gene specificity for steroid/thyroid hormone receptors. *Cell*, **57**, 1139–1146.

Umesono,K., Murakami,K.K., Thompson,C.C. and Evans,R.M. (1991) Direct repeats as selective response elements for the thyroid hormone, retinoic acid and vitamin D3 receptors. *Cell*, **65**, 1255–1266.

van Tilborg,M.A., Bonvin,A.M., Hard,K., Davis,A.L., Maler,B., Boelens,R., Yamamoto,K.R. and Kaptein,R. (1995) Structure refinement of the glucocorticoid receptor–DNA binding domain from NMR data by relaxation matrix calculations. *J. Mol. Biol.*, **247**, 689–700.

Vögtli,M., Elke,E., Imhof,M.O. and Lezzi,M. (1998) High level transactivation by the ecdysone receptor complex at the core recognition motif. *Nucleic Acids Res.*, **26**, 2407–2414.

Yao,T.P., Segraves,W.A., Oro,A.E., McKeown,M. and Evans,R.M. (1992) *Drosophila* ultraspiracle modulates ecdysone receptor function via heterodimer formation. *Cell*, **71**, 63–72.

Yao,T.P., Forman,B.M., Jiang,Z., Cherbas,L., Chen,J.D., McKeown,M., Cherbas,P. and Evans,R.M. (1993) Functional ecdysone receptor is the product of EcR and Ultraspiracle genes. *Nature*, **366**, 476–497.

Zechel,C., Shen,X.Q., Chambon,P. and Gronemeyer,H. (1994a) Dimerization interfaces formed between the DNA binding domains determine the cooperative binding of RXR/RAR and RXR/TR heterodimers to DR5 and DR4 elements. *EMBO J.*, **13**, 1414–1424.

Zechel,C., Shen,X.Q., Chen,J.Y., Chen,Z.P., Chambon,P. and Gronemeyer,H. (1994b) The dimerization interfaces formed between the DNA binding domains of RXR, RAR and TR determine the binding specificity and polarity of the full-length receptors to direct repeats. *EMBO J.*, **13**, 1425–1433.

Zhao,Q., Khorasanizadeh,S., Miyoshi,Y., Lazar,M.A. and Rastinejad,F. (1998) Structural elements of an orphan nuclear receptor–DNA complex. *Mol. Cell*, **1**, 849–861.

Zhao,Q., Chasse,S.A., Devarakonda,S., Sierk,M.L., Ahvazi,B. and Rastinejad,F. (2000) Structural basis of RXR–DNA interactions. *J. Mol. Biol.*, **296**, 509–520.

*Received July 9, 2003; revised August 27, 2003; accepted September 16, 2003*