Published online 10 April 2007

Nucleic Acids Research, 2007, Vol. 35, No. 8 ***2705–2718***
doi:10.1093/nar/gkm162

Novel DNA-binding element within the C-terminal extension of the nuclear receptor DNA-binding domain

Michał Jakób[1], Robert Kołodziejczyk[1], Marek Orłowski[1], Szymon Krzywda[2], Agnieszka Kowalska[1], Joanna Dutko-Gwóźdź[1], Tomasz Gwóźdź[1], Marian Kochman[1], Mariusz Jaskólski[2,3] and Andrzej Ożyhar[1,*]

[1]Department of Biochemistry, Faculty of Chemistry, Wrocław University of Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland, [2]Department of Crystallography, Faculty of Chemistry, A. Mickiewicz University, Poznań, Poland and [3]Center for Biocrystallographic Research, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznań, Poland

Received January 10, 2007; Revised March 2, 2007; Accepted March 5, 2007

ABSTRACT

The heterodimer of the ecdysone receptor (EcR) and ultraspiracle (Usp), members of the nuclear receptors superfamily, is considered as the functional receptor for ecdysteroids initiating molting and metamorphosis in insects. Here we report the 1.95 Å structure of the complex formed by the DNA-binding domains (DBDs) the EcR and the Usp, bound to the natural pseudopalindromic response element. Comparison of the structure with that obtained previously, using an idealized response element, shows how the EcRDBD, which has been previously reported to possess extraordinary flexibility, accommodates DNA-induced structural changes. Part of the C-terminal extension (CTE) of the EcRDBD folds into an α-helix whose location in the minor groove does not match any of the locations previously observed for nuclear receptors. Mutational analyses suggest that the α-helix is a component of EcR-box, a novel element indispensable for DNA-binding and located within the nuclear receptor CTE. This element seems to be a general feature of all known EcRs.

INTRODUCTION

Multicellular organisms require specific intercellular communication to properly orchestrate the complex body plan during embryogenesis and to maintain its function during the entire lifespan. Classical signal transduction cascades are initiated by ligand binding to membrane-anchored receptors, eventually changing the activity of specific nuclear transcription factors. In contrast, members of the nuclear receptor superfamily transduce their signals directly. The receptors have evolved to combine the functions of signal responsiveness, DNA-binding and transcriptional activation into one protein composed of functionally separated modules/domains (1). A typical nuclear receptor is composed of the N-terminal AB region, the DNA-binding C domain, a hinge D region, the ligand-binding E domain and the C-terminal F region. The core DNA-binding domain (DBD), which is a defining feature of the family (2), plays a central role in the correct positioning of the receptors, and complexes recruited by them, close to the genes whose transcription is affected (3). To achieve this aim, the DBD must overcome the challenge of finding small cognate response elements within the entire genome. A long-standing question is therefore, how this selection is achieved, given that nuclear receptors employ a highly conserved DBD and a set of response elements, which is quite limited in both sequence and structure. To solve this mystery on the molecular level, crystal structures of some nuclear receptor DBDs in complex with DNA have been analyzed. Unfortunately, most of the research in this field has been carried out using idealized, highly symmetric DNA duplexes, disregarding the fact that natural response elements are characterized by high-sequence degeneracy (3).

Representatives of the nuclear receptor superfamily have been identified in almost all classes of metazoans and the availability of complete genome sequences has revealed some interesting data regarding the occurrence of nuclear receptors. For example, the human genome sequence reveals 48 members of the family with 21 genes representing receptors with known ligands, usually small lipophilic molecules, including steroids, the thyroid

*To whom correspondence should be addressed. Tel: +48-71-320-6333; Fax: +48-71-320-6337; Email: andrzej.ozyhar@pwr.wroc.pl

hormone, retinoic acids and vitamin D (1). In contrast to the complexity of the human hormone signaling pathways, *Drosophila melanogaster* has only one lipophilic hormone acting as a nuclear receptor ligand, the steroid hormone 20-hydroxyecdysone (20E), and its genome contains only 18 nuclear receptor genes (4). This makes *D. melanogaster* an ideal system for studying nuclear-receptor function and regulation. The 20E hormone, which is considered to be a principal determinant of developmental timing, controls diverse biological processes, including morphogenetic, apoptotic, physiological, reproductive and behavioral responses (5,6). Like vertebrate steroid hormones, 20E exerts its effects via a member of the nuclear receptor superfamily, a product of the *EcR* gene (7). Although ecdysone receptor (EcR) can bind 20E on its own (8), the binding is greatly stimulated (9) by its heterodimerization partner, a product of the *ultraspiracle* gene (Usp), which is another member of the nuclear receptor superfamily and the *D. melanogaster* ortholog (10) of the mammalian retinoid X receptors (RXRs). Since it has been also observed that ligand binding stabilizes the EcR/Usp complex and increases its affinity for 20E-response elements, the EcR/Usp heterodimer is believed to be the only functional form of the 20E receptor (11). Although molecular studies of the EcR/Usp heterodimer are not as advanced as those of vertebrate heterodimeric receptors, it is already clear that the ecdysteroid receptor complex holds an exceptional position within the nuclear receptor family. It has been shown, for example, that the ligand-binding domain (LBD) of EcR is characterized by unusual flexibility and adaptability, which allows for molding of this domain around different ligands (12). Although 20E is thought to elicit most of the above-mentioned physiological responses, mounting evidence indicates, that alternate signaling pathways exist that are driven by ecdysteroids other than 20E, some of which are present at specific stages during development (13). According to a recent hypothesis, 'conformational compatibility' between the cognate receptor's LBD and an ecdysteroid molecule would determine the initiation of genomic versus non-genomic response pathways (14). Interestingly, the DBD of EcR appears to possess high plasticity as well (15). Therefore, EcR could adopt different, although ligand- and response-element-specific conformations evoking numerous ecdysteroid-dependent effects (14). Another feature distinguishing the EcR/Usp heterodimer from its vertebrate counterparts, which tend to form complexes on inherently asymmetric DNA-binding sites composed of directly repeated half-sites, is its propensity for response elements arranged as psudopalindromes with a single intervening nucleotide (16). Our mutational studies of the interaction of the Usp and EcR DBDs (UspDBD and EcRDBD, respectively) with the pseudopalindromic response element from the *hsp27* gene promoter (*hsp27pal*), have demonstrated that natural pseudopalindromic ecdysone response elements may act as functionally asymmetric elements that fix the Usp/EcR heterodimer in a specific orientation. In particular, it has been shown that UspDBD, which preferentially binds the 5′ half-site of *hsp27pal*, operates as a key factor dictating the polarity of the heterocomplex (5′-UspDBD-EcRDBD-3′) (16,17). Although this polarity was verified by the crystal structures of the UspDBD/EcRDBD and RXRDBD/EcRDBD (RXRDBD/EcRDBD) complexes with an idealized non-natural element organized as an inverted repeat of the 5′-AGGTCA-3′ sequence separated by 1 bp (IR-1) (18), many important observations coming from earlier biochemical studies could not be confirmed. In particular, the previously reported importance of the EcRDBD C-terminal extension (CTE) sequence for effective formation of the UspDBD/EcRDBD heterocomplexes (16) could not be explained because the crystallographic data allowed the visualization of only a few CTE residues, mostly in the RXRDBD/EcRDBD-IR-1 complex, but not in the UspDBD/EcRDBD-IR-1 complex.

To ultimately elucidate the molecular basis for DNA target specificity of the UspDBD/EcRDBD heterodimer, we have solved the crystal structure of both domains bound to the natural response element from the *hsp27* gene promoter. Our data reveal important molecular details of the UspDBD/EcRDBD–DNA interaction, which could not be observed previously, when unnatural DNA response element was used for crystallization. Most importantly, our crystallographic data demonstrate that part of the CTE of EcRDBD folds into an α-helix whose location in the minor groove does not match any of the locations previously observed for other CTEs of other nuclear receptors. Analysis of the crystallographic data, together with mutational analyses, suggest that the α-helix is a part of a novel DNA-binding supporting element, previously unobserved for any of the nuclear receptors. This element, which we refer to as EcR-box, seems to be a characteristic feature of all known EcRs.

MATERIALS AND METHODS

Construction of expression vectors, site-directed mutagenesis, overexpression and purification of the wild-type and mutant proteins

The plasmid pGEX-2T (Amersham Biosciences, Freiburg, Germany) containing the *lac*Iq gene was used for the expression of DBDs as fusion proteins with *Schistosoma japonicum* glutathione-S-transferase (GST) in *Escherichia coli* strain BL21(DE3)pLysS (Novagen, Germany). The construction of the expression plasmids for the wild-type *D. melanogaster* EcR, Usp GST-DBDs and the wild-type *Bombyx mori* EcR GST-DBD has been described previously (15,19). The expression plasmid for the wild-type *B. mori* UspDBD (pGEX-2T · $UspDBD_{(104–202)}$_Bm) was constructed using the following primers: 5′-gcccgg*ggatcc*GCACCTCGACAG CAAGAG-3′ (sense) and 5′-gcccgg*ggatcc*GTCTTCGAC TGTGGTCGTA-3′ (antisense). Small letters in the above sequences represent nucleotides added for cloning purpose whereas the restriction sites are shown in italics. The PCR-based megaprimer mutagenesis protocol (20) was used to introduce new alanine codons within the cDNA region encoding *D. melanogaster* EcRDBD A-box and standard PCR to introduce deletion mutations within the CTE of *D. melanogaster* and *B. mori* EcRDBD.

The plasmids: pGEX-2T · EcRDBD$_{(256–364)}$_Dm and pGEX-2T · EcRDBD$_{(199–307)}$_Bm (15) were used as templates. The sequences of the recombinant cDNA fragments were verified by dideoxy sequencing. Expression of GST-DBDs and purification of wild-type and mutated DBDs as GST-free proteins was performed using a procedure described previously for UspDBD with deleted C-terminal sequence (UspDBD$_{\Delta A/T}$) (16). The concentration of the purified proteins was determined spectrophotometrically at 280 nm using absorption coefficients calculated according to Gill and von Hippel (21).

DNA-binding assays

The electrophoretic mobility shift assays (EMSA) experiments (22) were performed under conditions described previously (15) using ^{32}P-labeled ds oligonucleotide 5′-AGCGACAAGGGTTCAATGCACTTGT CCAATGAA-3′ (only one strand is shown), based on the natural 20E pseudopalindromic response element (underlined) from the *D. melanogaster hsp27* gene promoter (23,24).

Crystallization and data collection

The protein–DNA complex was prepared using stoichiometric amounts of UspDBD, EcRDBD and duplex DNA (for details see Figure 1) concentrated to ~0.5 mM (each species) in 15 mM MES buffer, pH 5.5, 50 mM NaCl, 1 mM DTT, 5 μM $ZnCl_2$, using Amicon® Ultra-4 Centrifugal Filters (Millipore, Poland). Best crystals ($0.6 \times 0.1 \times 0.015$ mm) were grown at 20°C in sitting drops made of 1 μl of the protein–DNA complex solution and an equivalent volume of precipitant solution, containing 21% PEG 3350, 100 mM MES buffer, pH 5.5, 100 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT, 5 μM $ZnCl_2$, LiCl or 10 mM urea.

For X-ray diffraction experiments the crystal was flash-frozen to 100 K in nitrogen gas stream without additional cryoprotection. Diffraction data extending to 1.95 Å resolution were collected using a MAR Research CCD detector at the EMBL beamline X11 of the DESY synchrotron (Hamburg, Germany) operated at a wavelength $\lambda = 0.8115$ Å. The data were indexed, integrated and scaled using the HKL package (25). The crystals belong to the $P2_1$ space group with unit cell parameters $a = 46.7$, $b = 59.8$, $c = 65.2$ Å, $\beta = 106.72°$ and contain one complex molecule per asymmetric unit. Data collection statistics are summarized in Table 1.

Structure solution and refinement

The structure was solved by molecular replacement using the genetic algorithms EPMR program (26) and the UspDBD/EcRDBD-IR-1 structure (18) (PDB ID 1R0O) as a model. For the diffraction data between 15 and 4 Å, the program located the expected single copy of the complex in the asymmetric unit with a correlation coefficient of 55%.

The model was manually rebuilt into electron density maps during iterative cycles of modeling in Quanta 2000 (Accelrys Inc., San Diego, CA, USA) which alternated with maximum-likelihood restrained refinement as implemented in the Refmac5 program from the CCP4 suite (27). Two hundred and twenty-two water molecules were added using the X-Solvate module of Quanta 2000 at the final stages of the refinement. The final model contains 165 amino acid residues (78 UspDBD, 87 EcRDBD), 20 base-paired nucleotides, four zinc cations and 222 water molecules. The refinement converged with R and R_{free} factors of 0.180 and 0.217, respectively.

Coordinates

The atomic coordinates and structure factors have been deposited in the PDB with the accession code 2HAN.

RESULTS

Crystallization and structure determination

An 86-resiude fragment of *D. melanogaster* UspDBD consisting of residues −10 to +76 and a 109-residue fragment of *D. melanogaster* EcRDBD, residues −8 to +101 (Figure 1A, B), were expressed individually in *E. coli* and purified to homogeneity. The primary structures of the DBDs were the same as the structures of the DBDs used previously in crystallization studies with an idealized fully symmetric IR-1 element (18). Here, the purified DBDs were co-crystallized with a DNA fragment containing the natural pseudopalindromic 20E response element from the *D. melanogaster hsp27* gene promoter (*hsp27pal*, Figure 1C). The structure was solved by molecular replacement using the PDB coordinates 1R0O of the UspDBD/EcRDBD-IR-1 complex (18) as the search model. The asymmetric unit of the crystal contains one copy of the UspDBD/EcRDBD-DNA complex. The final model, refined to 1.95 Å resolution, contains residues −3 to +76 of UspDBD, residues −6 to +81 of EcRDBD, the complete DNA and 222 water molecules. The refinement is summarized in Table 1.

Four side chains of the protein components have clearly defined alternative conformations (see further). Three of them are located at crucial protein–protein and protein–nucleic acid interfaces. All the multiple conformations in the interface areas have unambiguous definition in electron density maps. The course of the refinement and the final model were validated using the R_{free} test (28).

Overall architecture of the complex

In agreement with biochemical data (16) and with the structure of UspDBD/EcRDBD bound to an idealized IR-1 element (18), the UspDBD/EcRDBD heterodimer is bound to the *hsp27pal* element with a defined polarity, where the UspDBD occupies the 5′-half-site of *hsp27pal* and EcRDBD the 3′-half-site. A superposition of the structure of the UspDBD/EcRDBD-*hsp27pal* complex with the UspDBD/EcRDBD-IR-1 structure is shown in Figure 2 and is characterized by an r.m.s.d.-value of 0.54 Å calculated for the corresponding Cα atoms. Although the overall fold of the two heterodimers is similar, the present structure of the UspDBD/EcRDBD complex interacting with the natural element reveals in

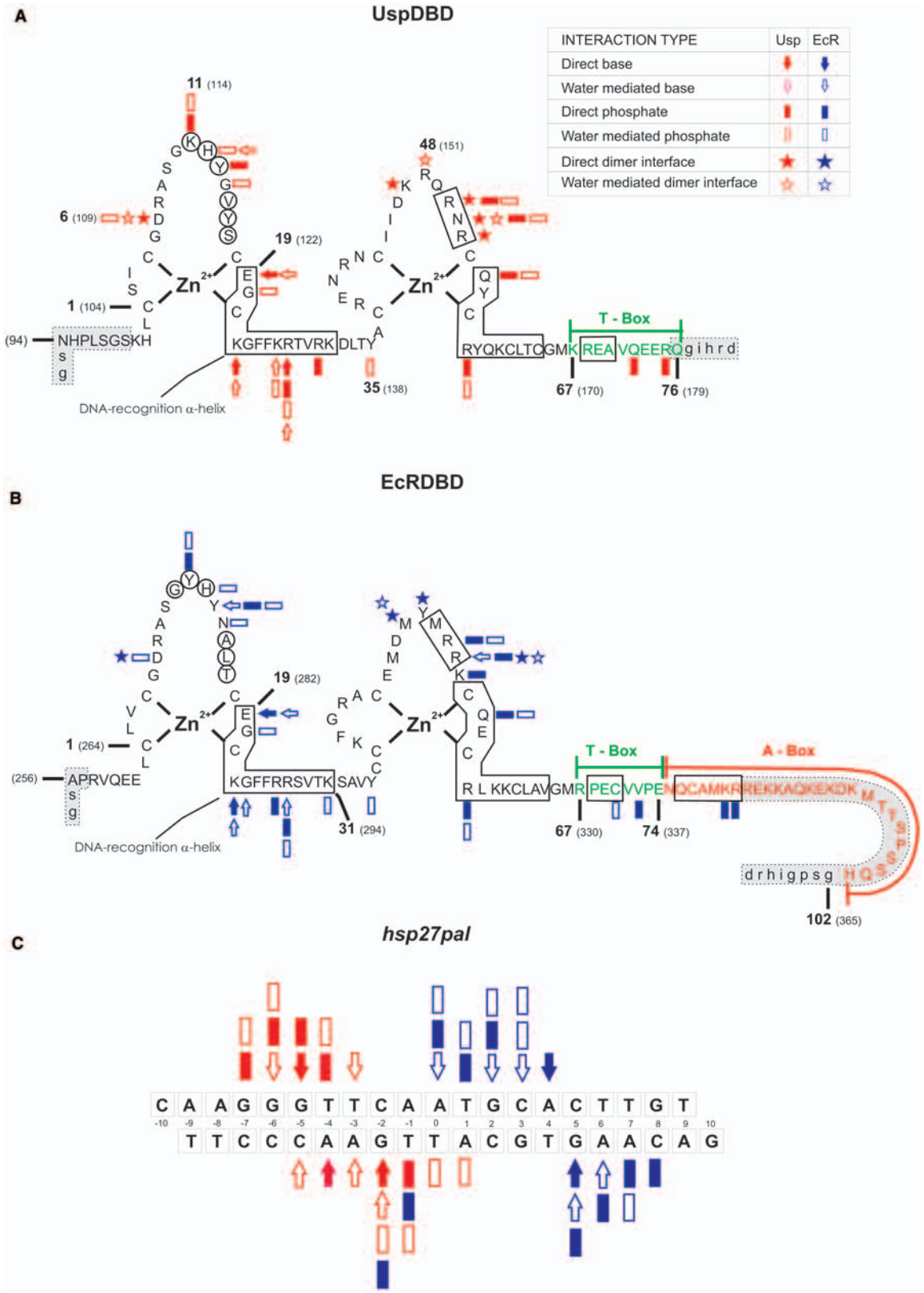

Figure 1. The protein and response element constructs used in crystallization and their contacts. Sequences and interactions (legend is provided within the figure) are shown for UspDBD (**A**) and EcRDBD (**B**), respectively. The numbering of the amino acid residues is relative to the first conserved cysteine, with the authentic numbers (7,10) appearing in parentheses The sequences defined previously (16) as corresponding to T-box (42) and A-box (33,34) are highlighted in green and red, respectively. In gray boxes the N- and C-terminal residues not visible in the electron density maps are listed. Cloning artifacts from the pGEX-2T plasmid are indicated by lower case letters. The α-helices are boxed and the residues from β-sheets are circled following the definition of DSSP (43). (**C**) The *hsp27pal*-based DNA used in cocrystallization. The symbols are as in (**A**) and (**B**).

Table 1. Crystallographic data. The values in parentheses represent the values for the last resolution shell.

Data collection	
Radiation source	Beamline X11, EMBL Hamburg
Wavelength (Å)	0.8115
Temperature (K)	100
Space group	$P2_1$
Unit-cell parameters (Å, °)	$a=46.7$, $b=59.8$, $c=65.2$, $\beta=106.7$
Resolution (Å)	30–1.95
Number of observed reflections	242 671
Number of unique reflections	25 368
Redundancy	9.6
Completeness (%)	99.9 (100)
R_{merge}	0.068 (0.549)
$<I/\sigma\ (I)>$	24.2 (3.4)
Refinement statistics	
Resolution range (Å)	26.92–1.95
Number of all/test reflections	23 899/1272
R/R_{free}	0.180/0.217
$<B>$ (Å^2)/number of atoms	
Protein	36.0/1348
DNA	40.9/812
Zn^{2+}	28.4/4
Solvent	38.1/222
R.m.s. deviations from ideal	
Bond lengths (Å)	0.016
Bond angles (°)	2.03
Ramachandran statistics (%)	
Most favored/additionally allowed regions	84.4/14.3

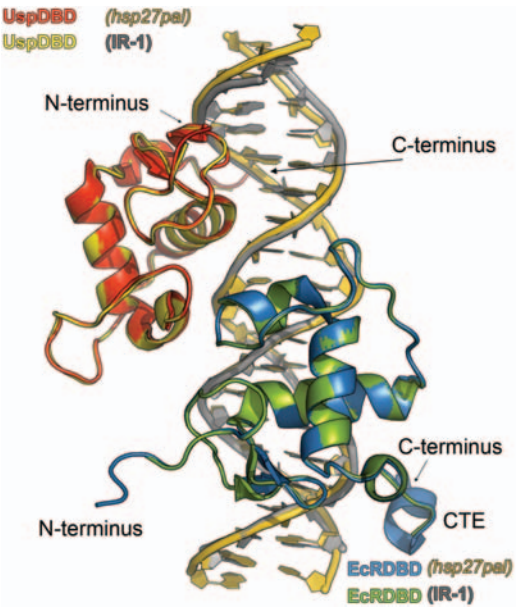

Figure 2. Superposition of the UspDBD/EcRDBD-*hsp27pal* and UspDBD/EcRDBD-IR-1 structures. Shown are UspDBD (red) and EcRDBD (blue) bound to *hsp27pal* oligonucleotide (gold), as well as UspDBD (yellow) and EcRDBD (green) bound to the IR-1 element (gray). The UspDBD/EcRDBD-IR-1 structure (18) is based on the coordinate file deposited in the PDB (accession code 1R0O).

total eleven additional amino acid residues: three at the N-terminus of EcRDBD (R-6, V-5, Q-4), six in the CTE of EcRDBD (Q76, C77, A78, M79, K80, R81), i.e., in the N-terminal region corresponding to the so-called A-box, and one residue at each end of UspDBD (K-3, R75) (Figure 1A, B). The N-terminal fragment of EcRDBD is directed away from the DNA molecule into the solvent and does not exhibit any specific interactions with the DNA molecule. A remarkable feature of the UspDBD/EcRDBD-*hsp27pal* complex is the presence of an α-helix at the C-terminus of the EcRDBD molecule, involved in interactions with the DNA. This fragment was disordered in EcRDBD complexed with UspDBD on the non-natural IR-1 element (18), and therefore its involvement in binding the DNA response element is reported here for the first time. Furthermore, this α-helix may be a part of a unique structural element indispensable for effective DNA-binding, previously unobserved for any of the nuclear receptors (see subsequently).

According to recently published fluorescence resonance energy transfer data, binding to the UspDBD/EcRDBD heterodimer induces a significant bend of the *hsp27pal* element. Steady-state data indicate a bend of about $23\pm3°$ whereas a value of $21\pm4°$ could be estimated using fluorescence decay measurements. These observations were reinforced by gel retardation experiments where the apparent bend was estimated as 20.9° (29). To examine if any distortion of the *hsp27pal* element could be observed in the crystal structure, we analyzed the coordinate file using the 3DNA software (30). The local helical parameters obtained from the 3DNA software were used as input to the Madbend program for calculation of the bend magnitude and global roll of the DNA molecules (31). Calculations using Madbend with a reference plane in the middle of the *hsp27pal* molecule indicated that this element is bent toward the minor groove by 20.9° (Figure 3A). Devarakonda *et al.* (18) reported that in the UspDBD/EcRDBD-IR-1 complex there was no significant distortion of the IR-1 element, except for the spacer. In a re-evaluation of those data, using the same 3DNA/Madbend procedure and the coordinate file deposited for the UspDBD/EcRDBD-IR-1 complexes in the PDB (accession code 1R0O) the bend angle could be calculated as 24.1° (Figure 3A). Therefore, we conclude that formation of UspDBD/EcRDBD complexes on different response elements is accompanied by well-defined deformation of the DNA architecture (Figure 3B).

The basis for the recognition of the *hsp27pal* sequence by the Usp and EcR core DBDs

As shown in Figure 3A the sequence of the naturally occurring *hsp27pal* element is highly degenerated. Five of the fifteen base pairs making up the two heptameric half-site sequences deviate from perfect palindromic sequence. In contrast, the idealized IR-1 is a fully symmetric palindrome (Figure 3A), which according to gel shift studies can bind the UspDBD/EcRDBD heterocomplex, and also the complex of the full-length Usp and EcR, with higher affinity than *hsp27pal* (16,32). A side-by-side comparison of the UspDBD/EcRDBD-*hsp27pal* structure solved here with the previously published crystallographic data of the UspDBD/EcRDBD-IR-1 complex (18), demonstrates that complexes with natural and idealized response elements differ significantly in the mode of protein–DNA and protein–protein interactions.

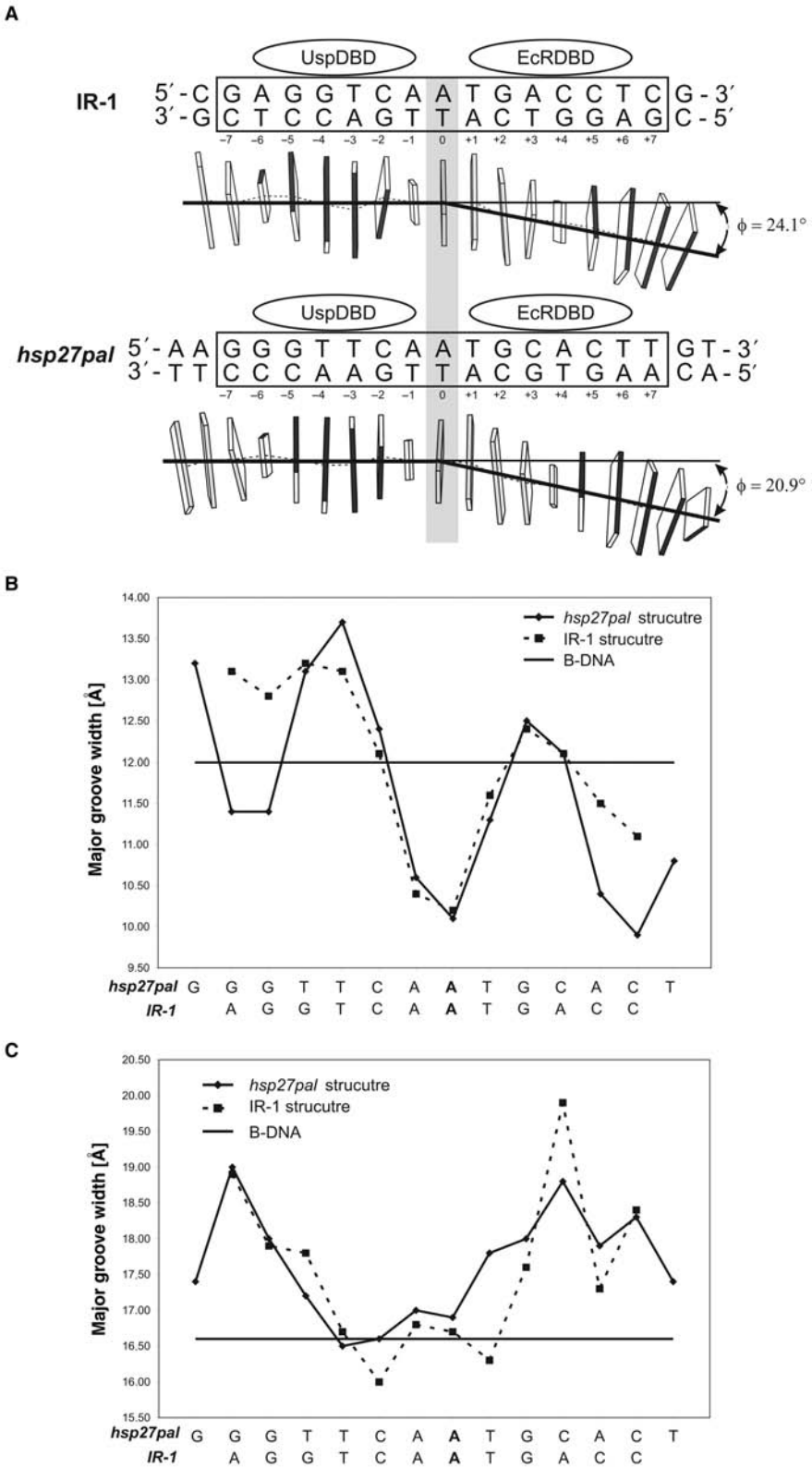

Figure 3. Evaluation of the crystallographic data indicates bending in *hsp27pal* and IR-1 elements. (**A**) The ds oligonucleotides used in the crystallographic analysis of the UspDBD/EcRDBD heterodimer complexed with the IR-1 and *hsp27pal* response elements. The sequences of the response elements are shown in frames, and ovals represent localization of the respective DBDs. Models and parameters of the DNA molecules were generated using the 3DNA software (30), data from this study and the atomic coordinate file of the UspDBD/EcRDBD-IR-1 structure (1R0O) (18). DNA base pairs are shown as rectangular blocks, and the idealized helical axis based on the axis computed by the 3DNA software (dotted line) is shown in black. (**B**) The minor and major grove widths of the *hsp27pal* (solid line) and IR-1 (broken line) ds oligonucleotides. Only the heptameric half-sites are shown. The values were derived using the 3DNA software. The solid black line represent standard values for B-DNA.

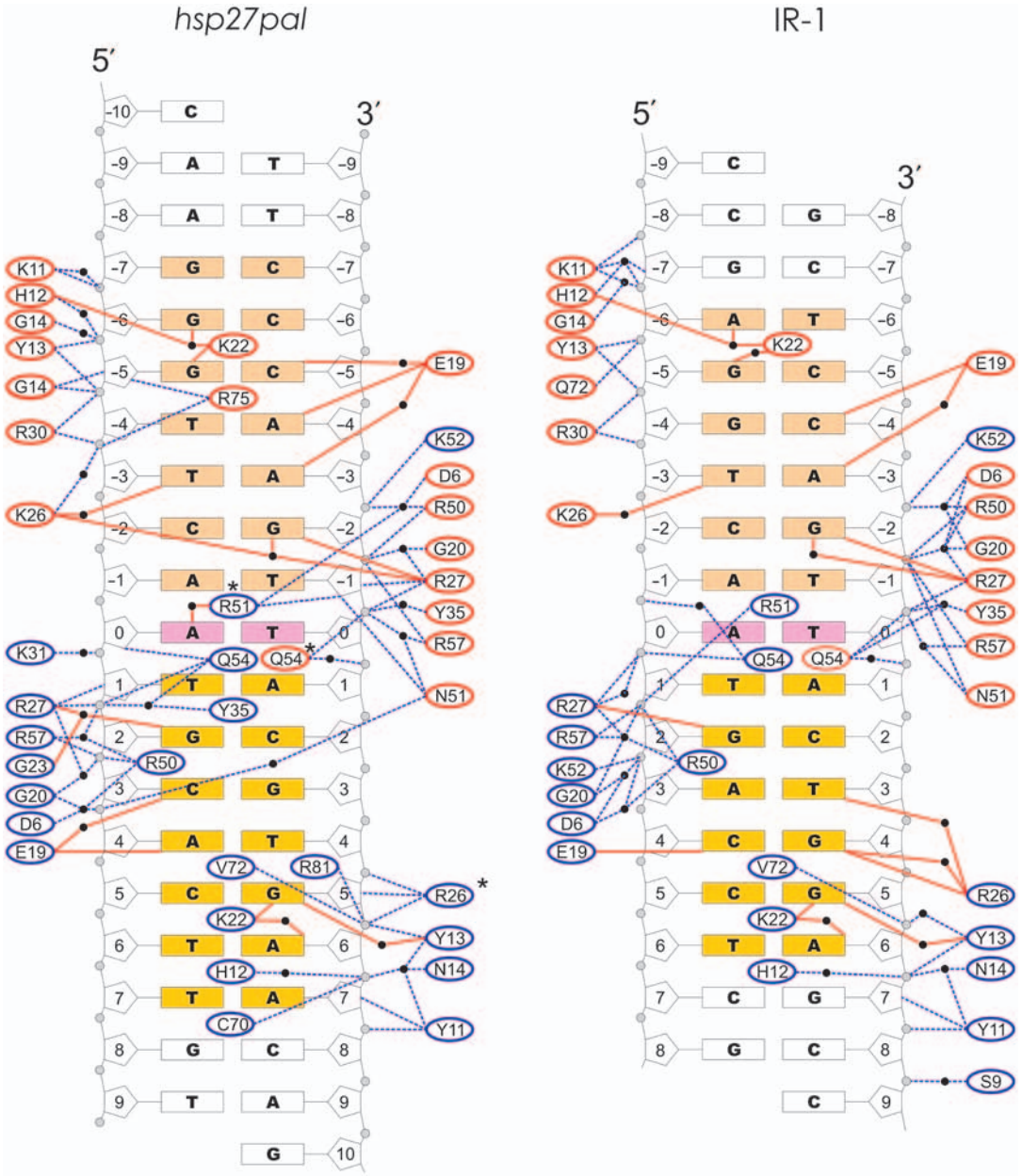

Figure 4. Comparison of protein–DNA contacts observed for the *hsp27pal* and IR-1 ds oligonucleotides bound by the UspDBD/EcRDBD heterocomplex. The crystallographic data obtained here for the UspDBD/EcRDBD-*hsp27pal* complex and published previously for the UspDBD/EcRDBD-IR-1 complex (PDB accession code 1R0O) (18) were analyzed in order to identify protein–DNA interactions. The DNA double helices are unwrapped to show schematically the base pairs and the respective contacts in the upstream (beige) and downstream (gold) half-sites of the *hsp27pal* and IR-1 elements. Gray circles represent phosphodiester groups, black circles are ordered water molecules that mediate protein-DNA contacts. Amino acid residues that contact the DNA using side chains are shown as red (EcRDBD) or blue (UspDBD) ovals. Asterisks denote residues which assume two well-defined conformations. Red solid lines indicate direct or water-mediated specific hydrogen bonding to base pairs. Blue dashed lines indicate hydrogen bonds to sugar-phosphate backbone.

The protein–DNA interactions for both complexes are depicted schematically in Figure 4. In agreement with previous structural characterizations of nuclear receptor DBD–DNA complexes (3), the key sequence-specific base contacts in the major groove are maintained in both complexes mostly by residues from the so-called DNA-recognition α-helix (for definition see Figure 1A, B). However, as shown in Figure 4, equivalent amino acid residues generate significantly different interaction patterns. Surprisingly, this is also true for residues interacting with identical DNA sequences, for example base pairs from positions −3 to −1. In total, ten interactions with the *hsp27pal* bases could be defined for the UspDBD molecule, three direct (involving E19, K22, R27) and seven water-mediated (involving K22, E19, K26, R27, H12). In comparison, for the UspDBD interacting with the IR-1 element only eight interactions with bases could be observed in total, two direct ones (E19, R27)

and six water-mediated (K22, E19, K26, R27, H12). EcRDBD on the other hand makes seven contacts with the *hsp27pal* 3′ half-site and eight with the 3′ half-site of the IR-1. However, the impact of direct or water-mediated contacts differs significantly for both elements. In particular, for the *hsp27pal* sequence there are two direct (K22, E19) and five water-mediated contacts (E19, K22, R27, R51) whereas for IR-1 four direct (E19, K22, R27, R26) and four water-mediated (Y13, K22, R26) contacts exist. Interestingly, only one residue, K22, of the EcRDBD exhibits the same sequence-specific interaction pattern in the UspDBD/EcRDBD-*hsp27pal* and UspDBD/EcRDBD-IR-1 structures.

A novel feature observed in the UspDBD/EcRDBD complex bound to the *hsp27pal* element is the ability of several amino acid residues of the EcRDBD to assume two well-defined conformations, which differ functionally. The first of these residues is R26 from the DNA-recognition α-helix. One of the R26 conformers makes contact with the phosphate backbone, whereas the other one, corresponding to that observed in EcRDBD complexed to the IR-1 element, is not involved in any interactions (Figure 5A, B). Similar observations could be made for Q54 (data not shown). Another residue assuming two conformations is R51, which belongs to a group of residues forming the subunit interface. Here, one of the conformers interacts with N51 of the UspDBD (Figure 5C), similarly as in the EcRDBD in the UspDBD/EcRDBD-IR-1 complex (not shown). Additionally, this conformer forms a novel *hsp27pal*-specific direct contact with the spacer A/T base-pair (Figure 5D). The second R51 conformer forms a set of direct and/or water-mediated contacts with UspDBD residues (Figure 5C, D). Dual conformation was also observed for R67, which is located on the surface of the EcRDBD molecule (not shown).

An α-helix of the CTE segment of EcRDBD is indispensable for efficient interaction with DNA

Earlier mutational and crystallographic studies of nuclear receptors, bound to asymmetric response elements organized as direct repeats, have emphasized the functional importance of the so-called CTE sequence of the core DBD. It has been shown that the CTE, consisting of the T-box and the adjacent α-helix (A-box), plays an important role in response element recognition, especially by vertebrate heterodimeric DBD complexes and by DBDs, which bind their cognate sequences as monomers (3). While the A-box residues are mainly involved in specific contacts with the response element, the T-box performs different functions, including specific base-pair recognition, formation of the dimer interface and support of the proper orientation of the A-box α-helix. In contrast to the high-sequence conservation of the core DBD region within the nuclear receptor superfamily, the CTE sequences are not preserved. Consequently it has been suggested that the CTE, as a DBD-characteristic element, would play an important role in response element discrimination by the DBDs interacting as hetero- and homodimers with the inherently asymmetrical directly repeated elements (33). The first experimental evidence suggesting the CTE would be also essential for the interaction of the DBDs with the palindromic response elements, was published by Niedziela-Majka *et al.* (16) who showed that deletion of the *D. melanogaster* EcRDBD CTE fragment encompassing the A-box sequence disrupted EcRDBD homo- and heterodimerization with UspDBD on *hsp27pal*. Moreover, detailed biochemical analyses demonstrated that the T-box of EcRDBD plays an important role in binding to *hsp27pal* (15). Unfortunately, these important functions of the CTE fragment could not be fully explained by the crystal structure of the UspDBD/EcRDBD heterocomplex bound to the idealized IR-1 element (18) where none of the A-box residues were included in the model. In contrast, in the present structure obtained for the natural *hsp27pal* element, seven residues of the EcRDBD A-box (N75-R81) are clearly seen in electron density. As shown in Figure 6, the residues form an α-helix, which definitely rests in the minor groove suggesting that this part of the EcRDBD may be involved in some interactions with the DNA. An analysis of the electron density suggests that the last two residues of the α-helix, K80 and R81, would be the prime candidates for such interactions (Figure 6). To validate this supposition we have obtained EcRDBD derivatives where K80 or R81 were substituted by alanine (Figure 7A). The binding affinities of the K80A and R81A mutants were determined by EMSA using a double-stranded oligonucleotide containing the original *hsp27pal* sequence. As *hsp27pal* had been shown previously to bind specifically and in a cooperative manner both the EcRDBD homo- and UspDBD/EcRDBD heterodimer, and since biochemical experiments indicate that A-box is necessary for these interactions (16), we tested the putative influence of the alanine substitutions on both homo- and heterodimer interactions. The binding of DNA by EcRDBD homodimers was clearly reduced by alanine substitutions at positions K80 and R81 (Figure 7B, C). The analyses presented in Figure 7D, E reveal that substitution of K80 and R81 by alanine is also detrimental to the effective formation of the UspDBD/EcRDBD-*hsp27pal* complex. Importantly, as demonstrated by EMSA results obtained for the respective alanine mutants (Figure 7A), this is also true for some other residues from the EcRDBD A-box, including K84, K85, Q87, K88 and K90 (Figure 7B–E), which are not visualized in the present structure. This observation is further supported by the EMSAs results obtained for EcRDBD CTE deletion mutants. As shown in Figure 8, deletion of the CTE up to the residue K92 does not change the affinity of the EcRDBD homo- and heterodimers (Figure 8B–E). In contrast, more extensive deletions encompassing residues identified here by means of crystallography and/or directed mutagenesis, reduced (deletion up to residue K85) or completely abolished (deletion up to N75) the binding of the homo- and/or heterodimers (Figure 8B, C and D, E). Together, the above results clearly indicate that some of the EcRDBD CTE residues, including K80 and R81, build up within the A-box a discrete functional entity, which is indispensable for the efficient interaction of the EcRDBD with DNA.

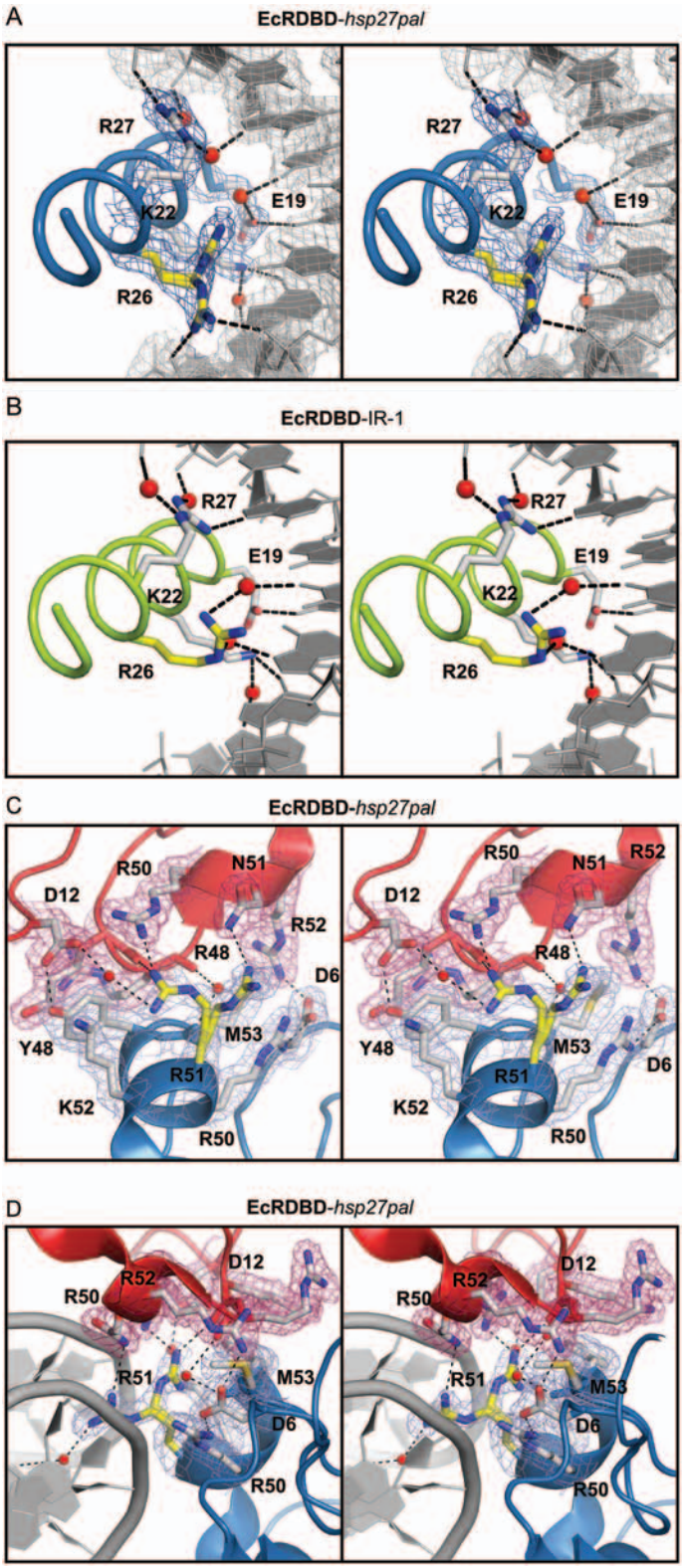

Figure 5. Four amino acids (R26, Q54, R51 and R67) of the EcRDBD assume two well-defined conformations in the UspDBD/EcRDBD complex bound to the *hsp27pal* element. The R26 residue is from the DNA-recognition α-helix. One of the R26 conformers makes contact with the phosphate backbone, whereas the second conformer, related to that observed in the IR-1 DNA complex, is not involved in any interactions (**A, B**). R51 is one of the residues forming the DBDs interface (**C, D**). One R51 conformer interacts with residue N51 of the UspDBD (**C**) and simultaneously forms a direct contact with the A/T base-pair (**D**). The second R51 conformer forms direct and water-mediated contacts with UspDBD (**C, D**). Hydrogen bonds in the stereodiagrams (A-D) are indicated as black dotted lines. Water molecules are shown as red spheres. For more details concerning other residues see text. The 2Fo-Fc electron density maps shown for selected side chains of the present UspDBD/EcRDBD-*hsp27pal* complex have been contoured at the 1.0 σ level.

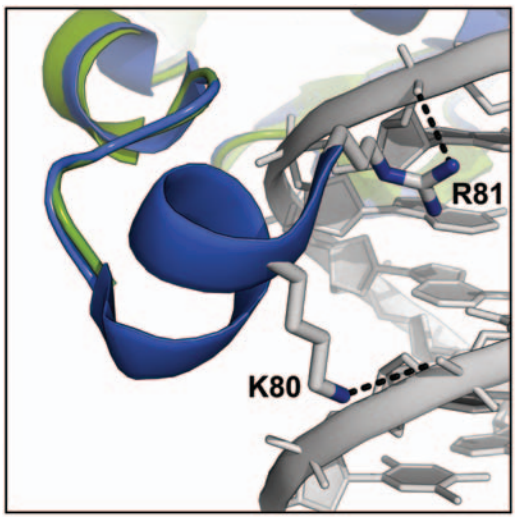

Figure 6. The A-box residues of the EcRDBD form an α-helical structure that interacts with the minor groove of the *hsp27pal* element. Comparison of the structures of the CTE fragments of the EcRDBD from the UspDBD/EcRDBD heterocomplex bound to the *hsp27pal* (blue) and IR-1 (green) elements. Dotted lines indicate hydrogen bonds formed by the K81 and R81 residues.

The mutational analyses presented in Figures 7 and 8 indicate that this conclusion is also true for the EcRDBD from *B. mori*.

DISCUSSION

The crystallographic data presented here for the UspDBD/EcRDBD heterodimer in complex with the natural pseudopalindromic response element from the *hsp27* gene promoter, and especially their comparison with the data published previously for the UspDBD/EcRDBD bound to the idealized IR-1 element (18), demonstrate the basis of the molecular rearrangements within both DBDs that permit them to adapt to different DNA sequences. Most importantly, the use of the natural element reveals some unexpected molecular details, which could not be observed in the structure published previously. Although the key sequence-specific base contacts are maintained in both structures essentially by the same amino acid residues from the DNA-recognition α-helices of the UspDBD and EcRDBD, the details of these interactions differ significantly and only one residue (K22 from EcRDBD) exhibits the same sequence-specific interaction pattern. At the same time, the UspDBD seems to form more sequence-specific contacts with the 5′ half of the *hsp27pal* than with the corresponding part of the IR-1. A total of ten and eight interactions could be identified, respectively. On the other hand, EcRDBD forms seven contacts with the *hsp27pal* sequence and eight with the IR-1 element. Furthermore, the impact of water-mediated contacts is significantly higher when the EcRDBD interacts with the natural element. In particular, five of seven contacts belong to this category whereas water mediates four of eight contacts with the IR-1 element. We emphasize that these significant *hsp27pal*-specific changes in the interaction pattern uncover unusual, previously not observed, molecular characteristics of the EcRDBD, such as dual conformation of the side chains of four amino acid residues. Only one of them, R67, situated at the surface of the domain, is apparently not involved in any protein–protein or protein–DNA interactions. The rotamers of two residues (R26 and Q54) are involved in *hsp27pal*-specific protein–DNA interactions. Finally, the R51 residue, located at the subunit interface, uses two conformations to create the direct and the water-mediated contacts with the UspDBD and with the *hsp27pal* sequence. We speculate that the alternate side chain conformations indicate that the EcRDBD molecule, functioning as part of the heterocomplex bound to the natural element but not to the unnatural symmetric element, retains some extraordinary structural flexibility reported previously for the isolated (i.e. not interacting with DNA and with UspDBD) domain (15). It has been also suggested that due to this property EcRDBD could easily accommodate DNA-induced changes in the secondary and tertiary structure (15). Indeed, part of the CTE of the EcRDBD bound in complex with the UspDBD on the *hsp27pal* folds into a novel α-helix not observed in the structure of the UspDBD/EcRDBD-IR-1 complex (18). The EMSA experiments presented here clearly demonstrate that the α-helix seems to be a component of a well-defined functional element, which is absolutely necessary for the effective formation of the UspDBD/EcRDBD-*hsp27pal* complex. The element extends from N75 to K92 within the CTE part previously defined (16) as corresponding to the so-called A-box (34). As indicated by the previously published alignment of the EcRDBD sequences (15), amino acid residues of the EcR CTE fragment exhibit a remarkable conservation although this fragment is not present in other nuclear receptors. As noted before, the sequence of *B. mori* EcR exhibits some puzzling differences (15). Nevertheless, as demonstrated by the EMSA results presented here, the CTE of the *B. mori* contains a well-defined fragment (P75-G89), which is critical for the formation of the UspDBD/EcRDBD-*hsp27pal* complex as well. Thus, the presence within the CTE of an element supporting DNA-binding, which we refer to as the EcR-box, seems to be a general feature of the EcR proteins. Taking into account our crystallographic data, which show that at least the N-terminal part of the *D. melanogaster* EcR-box (residues N75–R81) forms an α-helix, as well as secondary structure predictions done for the *D. melanogaster* and *B. mori* EcRDBDs (Figure 9A), we speculate that the entire EcR-box could fold into an α-helix (see subsequently).

As discussed earlier, in addition to the core DBD, the variable CTE region has been implicated in DNA response recognition and discrimination by the particular receptor. In contrast to the core DBD, which has the same structure in all receptor DBDs solved to date (3), the structure of each CTE fragment determined thus far has been unique. Moreover, distinct functions have been ascribed to the established CTEs. The crystal structure of the RXR/thyroid receptor DBDs heterodimer (RXRDBD/TRDBD) on its response element was the first to reveal the α-helical structure of the TR CTE.

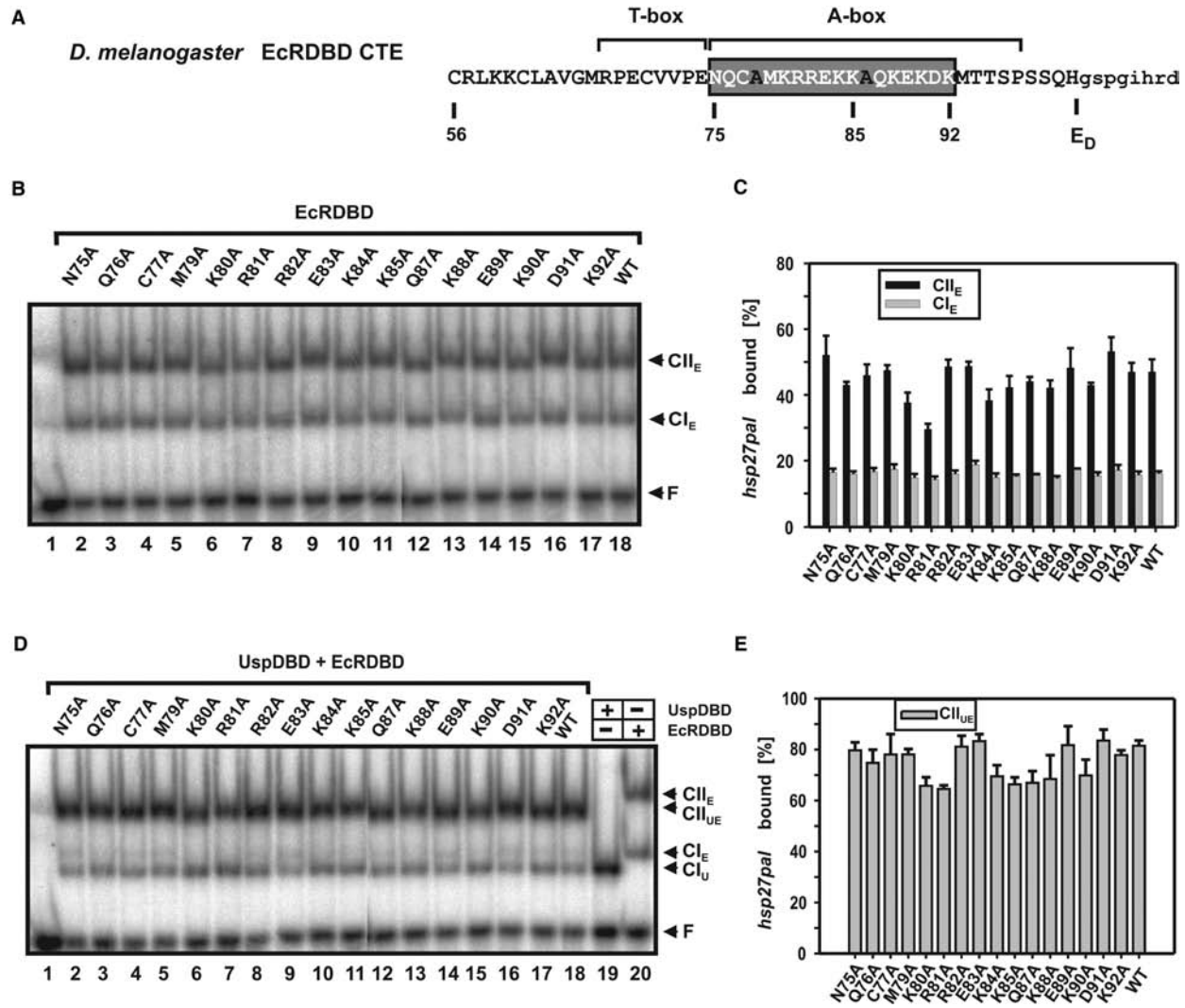

Figure 7. Effects of amino acids substitutions in the A-box of *D. melanogaster* EcRDBD on the interaction with the *hsp27pal* element. Sixteen individual residues from the A-box of the EcRDBD (N75–K92) (16) (**A**) were substituted with alanine. EMSA experiments were carried out with an *hsp27pal*-containing ds oligonucleotide and with the indicated homogenous EcRDBD (**B**) or equimolar mixture of the respective EcRDBD and the wild-type UspDBD (**D**). Panels (**C**) and (**E**) represent quantitative analysis of the EMSA data presented in panel (**B**) (lanes 2–18) and panel (**D**) (lanes 2–18), respectively. The *columns* indicate mean values of three independent experiments and error *bars* indicate SD-values. The designations of the respective mutant EcRDBDs are based on the amino acid single-letter code. The respective complexes formed by one DBD molecule are indicated by CI, and those originating from homo- or heterodimers are indicated by CII; F, free probe. For clarity, the wild-type complexes formed by EcRDBD are denoted as E and by UspDBD as U. The protein concentrations were: (**B**), lanes 2–18, 240 nM of the indicated EcRDBD; lane 19, the same amount of the wild-type UspDBD; lane 20, 120 nM of each wild-type DBD; (**E**), lanes 2–18, 120 nM of wild-type UspDBD and 120 nM of the indicated EcRDBD; lane 19, 240 nM of wild-type UspDBD; lane 20, 240 nM of wild-type EcRDBD.

It makes no tertiary contacts with the rest of the DBD but instead projects across the minor grove, where it makes an extensive interface with DNA (33) (Figure 9B). Although the CTE of the vitamin D receptor (VDR) DBD bears striking structural resemblance to the CTE of the TR (Figure 9B), it makes quantitatively different interactions with its cognate response element. It has been suggested that the primary role of the VDR CTE is to mediate response element discrimination, and not to provide additional DNA affinity (35). For other nuclear receptor DBDs, including nerve growth factor-inducible factor B (NGFI-B) (36), human liver receptor homologue-1 (LRH-1) (37) and RevErbα (38) where the CTEs do not fold into a defined α-helix (Figure 9B), it has been observed that they trace over one of the phosphate backbones and descend into the major groove contacting base pairs located 5′ to the response element half-site. The EcR-box fragment observed here on the *hsp27pal* element appears to have a novel and unique structural and possibly also multifunctional characteristic. First, being an α-helix the fragment does not project across the minor groove, as it was observed for TR and VDR CTEs, but descends into the minor groove, similarly as it was observed for non-helical CTEs (Figure 9B). However, due to the γ-turn formed by the P73, E74 and N75 residues, the orientation of this fragment, and possibly of the entire EcR-box, does

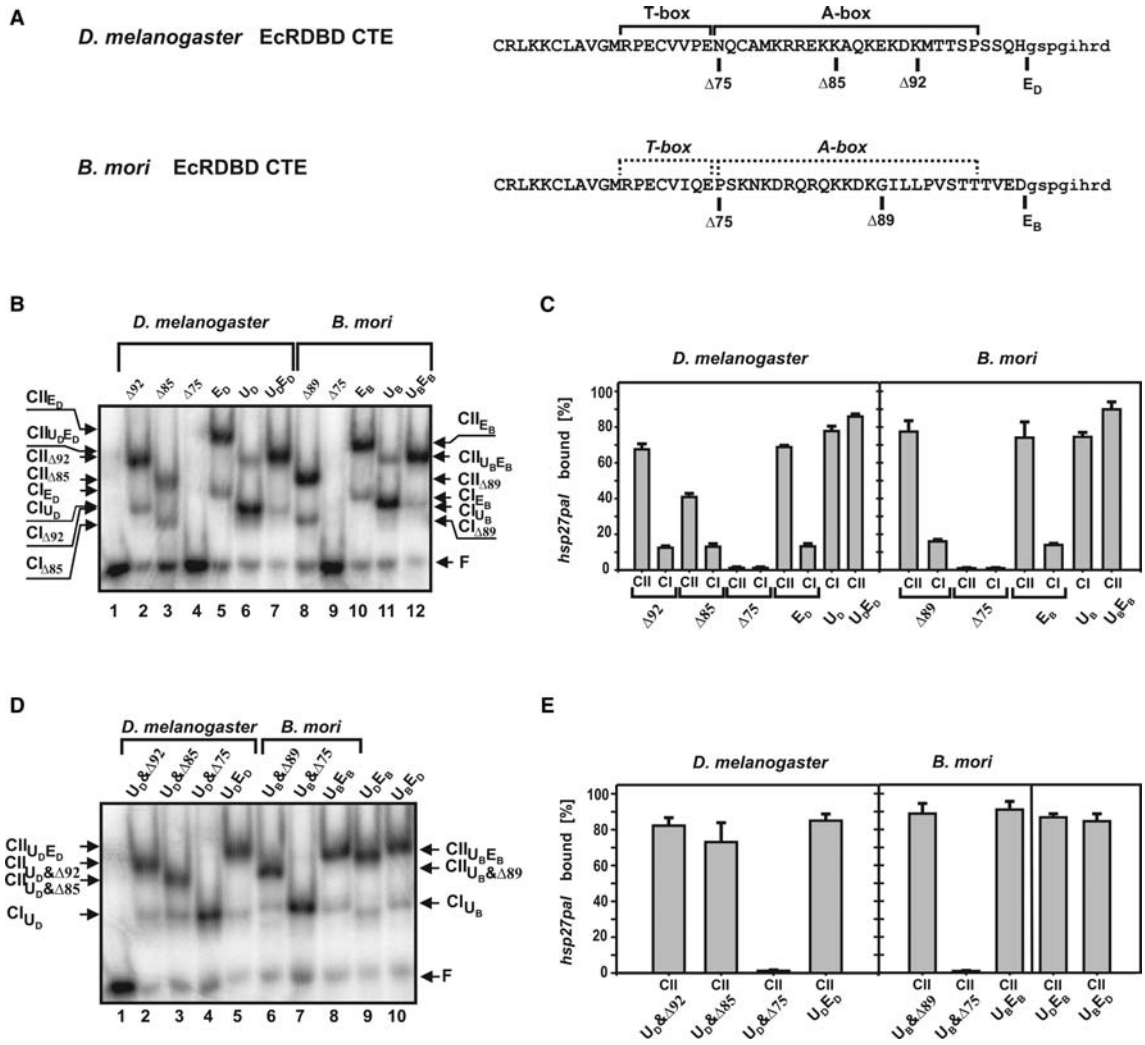

Figure 8. DNA-binding activities of CTE-truncated EcRDBDs of *D. melanogaster* and *B. mori*. Schematic diagram of the primary structure of the wild-type *D. melanogaster* and *B. mori* EcRDBD CTEs and their truncated derivatives, in which fragments from the region corresponding to the A-box (15,16) have been deleted (**A**). The numbering is relative to the first Zn^{2+}-coordinating cysteine of the DBDs. In order to analyze the effect of the deletion on the EcRDBD activity, the EMSA experiments were carried out with an *hsp27pal*-containing ds oligonuleotide and with the indicated EcRDBD derivative or wild-type EcRDBD (**B**), or with an equimolar mixture of the respective EcRDBD and wild-type UspDBD (**D**). Panels (**C**) and (**E**) represent quantitative analysis of the EMSA data presented in panel B (lanes 2–12) and panel D (lanes 2–10), respectively. For clarity, the wild-type full-length EcRDBDs were denoted by E_D for EcRDBD from *D. melanogaster* and by E_B for EcRDBD from *B. mori*. Similarly, the full-length UspDBDs from *D. melanogaster* and *B. mori* were designated as U_D and U_B, respectively. Other details are as in the legend to Figure 7. The protein concentrations were: (**B**), lanes 2–5 and 8–10, 600 nM of the indicated EcRDBD; lanes 6 and 11, 600 nM of the indicated wild-type UspDBD; lanes 7 and 12, 300 nM of each indicated wild-type DBD; (**D**), lanes 2–10, 300 nM of the indicated wild-type UspDBD and 300 nM of the indicated EcRDBD.

not match any of the previously observed orientations of the CTEs (Figure 9B, C), including that recently reported (39) for the homodimeric complex of the progesterone receptor DBDs (not shown). Our crystallographic data along with the EMSA experiments suggest that this orientation of the EcR-box would allow interactions of some amino acid residues in the minor groove. For the other above-mentioned proteins, which insert the CTE into the minor groove, numerous hydrogen-bonding contacts, including sequence-specific base contacts, have been observed. Our crystallographic data allow the identification of two residues involved in contacts with the sugar-phosphate backbone. Other key residues, indicated by the EMSA experiments, could not be modeled in the electron density maps, possibly indicating that they are located in an EcR-box fragment with increased dynamic properties. Apparently, interaction with the *hsp27pal* element is not sufficient to bring about a disorder-to-order transition of this EcR-box section and other factors are needed to achieve this. According to recent suggestions the predominant molecular function of disordered protein segments appears to involve molecular recognition in eukaryotes. The flexibility of the disordered fragments would allow them to be targets for multiple binding partners and post-translational modifications and thus enhance their ability to participate in multiple

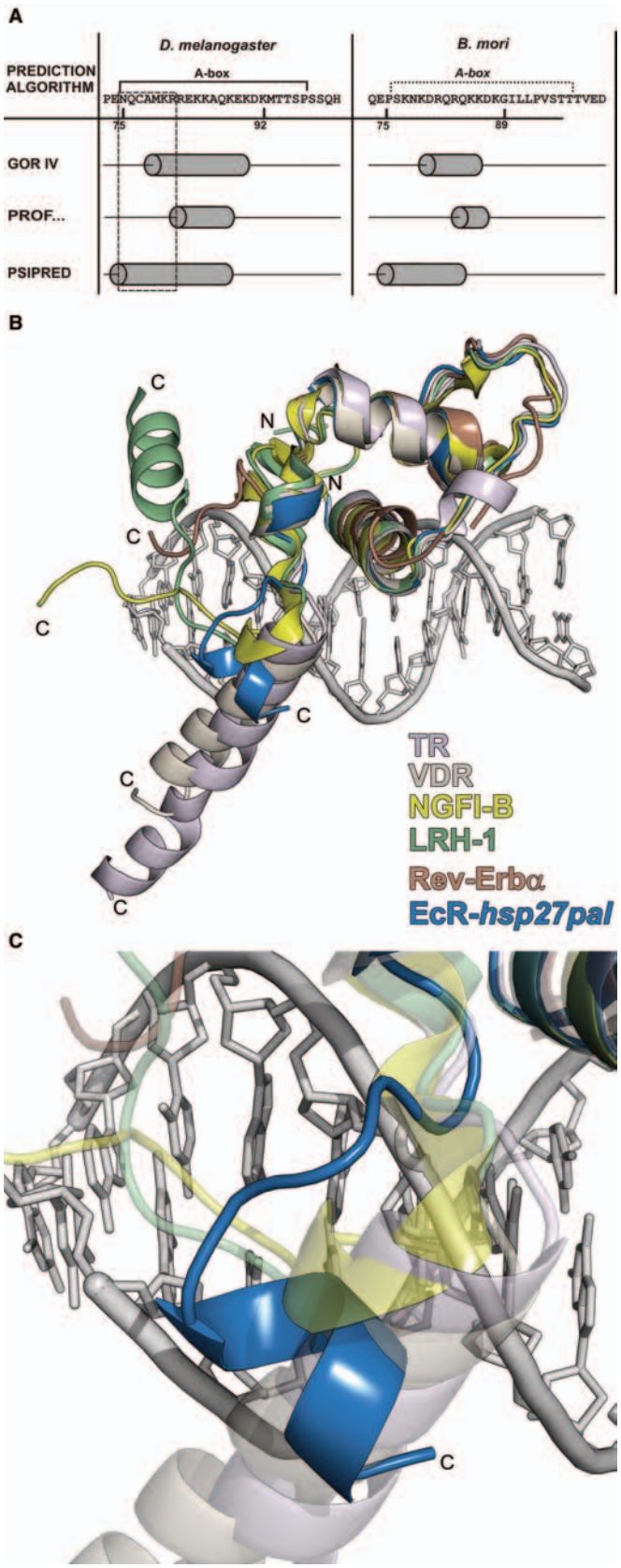

Figure 9. CTE of the EcRDBD folds into an α-helix which is part of the EcR-specific element supporting DNA-binding. (**A**) Secondary structure predictions of *D. melanogaster* and *B. mori* A-box segments. The predictions were performed using the GORIV, PROF and PSIPRED algorithms available at the ExPASy Proteomics Server (http://au.expasy.org/). The predicted α-helices are depicted as cylinders. The seven-residue segment forming the short α-helix in the CTE A-box of the *D. melanogaster* EcRDBD, observed in the present crystal structure, is boxed with broken line. The average confidence score values of the GORIV algorithm for the residues forming a potential α-helix within the *D. melanogaster* and *B. mori* A-box segments are 736 and 661, respectively. The PROF algorithm gives a prediction and a confidence value between 0 and 1 for each position in the amino acid sequence. The average confidence values for the *D. melanogaster* and *B. mori* sequences are 0.68 and 0.53, respectively. The PSIPRED algorithm gives a confidence value between 0 and 9. The average confidence scores for the *D. melanogaster* and *B. mori* are 5.1 and 4.8, respectively. (**B**) Superposition of the corresponding Cα atoms of the TR (PDB accession code 2NLL) (33), VDR (PDB accession code 1KB4) (35), LRH-1 (PDB accession code 2A66) (37), Rev-ERBα (PDB accession code 1A6Y) (38), NGFI-B (PDB accession code 1CIT) (36) and EcRDBD (from the present UspDBD/EcRDBD-*hsp27pal* complex) proteins complexed to DNA. (**C**) A detailed view of the CTEs and DNA shown in (**B**). C and N denote C- and N-termini of respective DBDs.

signaling pathways (40,41). For some nuclear receptors it has been demonstrated that the CTE can serve as a recruitment site for co-activators and as a target for post-translational modification (3). Although the molecular mechanism linking these regulatory events with control of the activity of the receptors remains elusive, it is reasonable to assume that also the EcR-box could be involved in similar processes. However, the questions of possible EcR-box interaction partners and of its post-translational modifications are still open.

SUPPLEMENTARY DATA

Supplementary Data is available at NAR Online.

ACKNOWLEDGEMENTS

We are grateful to Professor Jacek Otlewski (Institute of Biochemistry and Molecular Biology, University of Wrocław) for giving us the opportunity to perform crystallization experiments. We thank Professor Kostas Iatrou and Dr Luc Swevers (Institute of Biology, National Center for Scientific Research 'Demokritos,' Athens, Greece) for the kind gift of the cDNA clone encoding the full length of *B. mori* EcR, Dr George Tzertzinis (New England Biolabs, Beverly, MA, USA) for plasmid encoding the full length of *B. mori* Usp. In addition, we thank Professor Olaf Pongs and Dr Dirk Isbrandt (Center for Molecular Neurobiology, Hamburg, Germany) for their generous support. Supported by a grant number 3 T09A 040 28 from the State Committee for Scientific Research and by a subsidy from the Foundation for Polish Science to M.J. Some of the calculations were carried out in the Poznań Metropolitan Supercomputing and Networking Center. Funding to pay the Open Access publication charges for this article was provided by the State Committee for Scientific Research (grant number 3 T09A 040 28).

Conflict of interest statement. None declared.

REFERENCES

1. Escriva,H., Bertrand,S. and Laudet,V. (2004) The evolution of the nuclear receptor superfamily. *Essays Biochem.*, **40**, 11–26.
2. Evans,R.M. (1988) The steroid and thyroid hormone receptor superfamily. *Science*, **240**, 889–895.
3. Claessens,F. and Gewirth,D.T. (2004) DNA recognition by nuclear receptors. *Essays Biochem.*, **40**, 59–72.
4. King-Jones,K. and Thummel,C.S. (2005) Nuclear receptors—a perspective from Drosophila. *Nat. Rev. Genet.*, **6**, 311–323.
5. Riddiford,L.M., Cherbas,P. and Truman,J.W. (2000) Ecdysone receptors and their biological actions. *Vitam. Horm.*, **60**, 1–73.
6. Spindler-Barth,M. and Spindler,K.-D. (2000). In: Dorn,A (ed), *Progress in Developmental Endocrinology*. Wiley-Liss, New York, pp. 117–144.
7. 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.
8. Grebe,M., Fauth,T. and Spindler-Barth,M. (2004) Dynamic of ligand binding to Drosophila melanogaster ecdysteroid receptor. *Insect Biochem. Mol. Biol.*, **34**, 981–989.
9. 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.
10. Oro,A.E., McKeown,M. and Evans,R.M. (1992) The Drosophila retinoid X receptor homolog ultraspiracle functions in both female reproduction and eye morphogenesis. *Development*, **115**, 449–462.
11. 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**, 476–479.
12. Billas,I.M., Iwema,T., Garnier,J.M., Mitschler,A., Rochel,N. and Moras,D. (2003) Structural adaptability in the ligand-binding pocket of the ecdysone hormone receptor. *Nature*, **426**, 91–96.
13. Gilbert,L.I., Rybczynski,R. and Warren,J.T. (2002) Control and biochemical nature of the ecdysteroidogenic pathway. *Annu. Rev. Entomol.*, **47**, 883–916.
14. Schlattner,U., Vafopoulou,X., Steel,C.G., Hormann,R.E. and Lezzi,M. (2006) Non-genomic ecdysone effects and the invertebrate nuclear steroid hormone receptor EcR—new role for an 'old' receptor? *Mol. Cell. Endocrinol.*, **247**, 64–72.
15. Orłowski,M., Szyszka,M., Kowalska,A., Grad,I., Zoglowek,A., Rymarczyk,G., Dobryszycki,P., Krowarsch,D., Rastinejad,F., Kochman,M. *et al.* (2004) Plasticity of the ecdysone receptor DNA binding domain. *Mol. Endocrinol.*, **18**, 2166–2184.
16. 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.
17. 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.
18. Devarakonda,S., Harp,J.M., Kim,Y., Ożyhar,A. and Rastinejad,F. (2003) Structure of the heterodimeric ecdysone receptor DNA-binding complex. *EMBO J.*, **22**, 5827–5840.
19. Niedziela-Majka,A., Rymarczyk,G., Kochman,M. and Ożyhar,A. (1998) GST-Induced dimerization of DNA-binding domains alters characteristics of their interaction with DNA. *Protein Expr. Purif.*, **14**, 208–220.
20. Barik,S. (1995) Site-directed mutagenesis by double polymerase chain reaction. *Mol. Biotechnol.*, **3**, 1–7.
21. Gill,S.C. and von Hippel,P.H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.*, **182**, 319–326.
22. Fried,M. and Crothers,D.M. (1981) Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucleic Acids Res.*, **9**, 6505–6525.
23. Riddihough,G. and Pelham,H.R. (1987) An ecdysone response element in the Drosophila hsp27 promoter. *EMBO J.*, **6**, 3729–3734.
24. Ożyhar,A. and Kiltz,H.H. (1991) High-resolution gel filtration of the ecdysteroid receptor-DNA complex—an alternative to the electrophoretic mobility shift assay. *J. Chromatogr.*, **587**, 11–17.
25. Otwinowski,Z. and Minor,W. (1997). In: Carter,C and Sweet,R (eds), *Methods in Enzymology*, vol. 276. Academic Press, New York, pp. 307–326.
26. Kissinger,C.R., Gehlhaar,D.K. and Fogel,D.B. (1999) Rapid automated molecular replacement by evolutionary search. *Acta Crystallogr. D Biol. Crystallogr.*, **55**, 484–491.
27. Murshudov,G.N., Vagin,A.A. and Dodson,E.J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.*, **53**, 240–255.
28. Brunger,A. (1992) Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature*, **355**, 472–475.
29. Dobryszycki,P., Grad,I., Krusiński,T., Michaluk,P., Sawicka,D., Kowalska,A., Orłowski,M., Jakób,M., Rymarczyk,G., Kochman,M. *et al.* (2006) The DNA-binding domain of the ultraspiracle drives deformation of the response element whereas the DNA-binding domain of the ecdysone receptor is responsible for a slight additional change of the preformed structure. *Biochemistry*, **45**, 668–675.
30. Lu,X.J. and Olson,W.K. (2003) 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res.*, **31**, 5108–5121.
31. Strahs,D. and Schlick,T. (2000) A-Tract bending: insights into experimental structures by computational models. *J. Mol. Biol.*, **301**, 643–663.
32. Vogtli,M., Elke,C., 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.
33. 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.
34. Wilson,T.E., Paulsen,R.E., Padgett,K.A. and Milbrandt,J. (1992) Participation of non-zinc finger residues in DNA binding by two nuclear orphan receptors. *Science*, **256**, 107–110.
35. Shaffer,P.L. and Gewirth,D.T. (2002) Structural basis of VDR-DNA interactions on direct repeat response elements. *EMBO J.*, **21**, 2242–2252.
36. Meinke,G. and Sigler,P.B. (1999) DNA-binding mechanism of the monomeric orphan nuclear receptor NGFI-B. *Nat. Struct. Biol.*, **6**, 471–477.
37. Solomon,I.H., Hager,J.M., Safi,R., McDonnell,D.P., Redinbo,M.R. and Ortlund,E.A. (2005) Crystal structure of the human LRH-1 DBD-DNA complex reveals Ftz-F1 domain positioning is required for receptor activity. *J. Mol. Biol.*, **354**, 1091–1102.
38. 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.
39. Roemer,S.C., Donham,D.C., Sherman,L., Pon,V.H., Edwards,D.P. and Churchill,M.E. (2006) Structure of the progesterone receptor-DNA complex: novel interactions required for binding to half-site response elements. *Mol. Endocrinol.*, **12**, 3042–3052.
40. Tompa,P. (2005) The interplay between structure and function in intrinsically unstructured proteins. *FEBS Lett.*, **579**, 3346–3354.
41. Fink,A.L. (2005) Natively unfolded proteins. *Curr. Opin. Struct. Biol.*, **15**, 35–41.
42. Lee,M.S., Kliewer,S.A., Provencal,J., Wright,P.E. and Evans,R.M. (1993) Structure of the retinoid X receptor alpha DNA binding domain: a helix required for homodimeric DNA binding. *Science*, **260**, 1117–1121.
43. Laskowski,R.A., MacArthur,M.W., Moss,D.S. and Thornton,J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.*, **26**, 283–291.