Unique functional properties of a member of the Fushi Tarazu-Factor 1 family from *Schistosoma mansoni*

Benjamin BERTIN*, Souphatta SASORITH†, Stéphanie CABY*, Frédéric OGER*, Jocelyne CORNETTE*, Jean-Marie WURTZ† and Raymond J. PIERCE*¹

*INSERM U547, Institut Pasteur de Lille, 1 rue du Professeur Calmette, 59019-Lille, France, and †Département de Biologie et Génétique Structurales, Institut de Génétique et de Biologie Moléculaire et Cellulaire, 1 rue Laurent Fries, B.P. 163, 67404-Ilkirch, France

SmFtz-F1 (*Schistosoma mansoni* Fushi Tarazu-Factor 1) belongs to the Ftz-F1 subfamily of nuclear receptors, but displays marked structural differences compared with its mammalian homologues SF-1 (steroidogenic factor-1) or liver receptor homologue-1. These include a long F domain (104 amino acids), an unusually large hinge region (133 amino acids) and a poorly conserved E-domain. Here, using Gal4 constructs and a mammalian two-hybrid assay, we have characterized the roles of these specific regions both in the transcriptional activity of the receptor and in its interactions with cofactors. Our results have shown that, although the AF-2 (activation function-2) region is the major activation function of the receptor, both the F and D domains are essential for AF-2-dependent activity. Modelling of SmFtz-F1 LBD (ligand-binding domain) and structure-guided mutagenesis allowed us to show the important role of helix H1 in maintaining the structural conformation of the LBD, and suggested that its autonomous transactivation activity, also observed with SF-1, is for-

tuitous. This strategy also allowed us to study an eventual ligand dependence for this orphan receptor, the predicted three-dimensional models suggesting that the SmFtz-F1 LBD contains a large and well-defined ligand-binding pocket sealed by two arginine residues orientated towards the interior of the cavity. Mutation of these two residues provoked a loss of transcriptional activity of the receptor, and strongly reduced its interaction with SRC1 (steroid receptor cofactor-1), suggesting a ligand-dependent activity for SmFtz-F1. Taken together, our results argue for original and specific functional activities for this platyhelminth nuclear receptor.

Key words: Fushi Tarazu-Factor 1 (Ftz-F1), Gal4 fusion protein, nuclear hormone receptor, *Schistosoma mansoni*, site-directed mutagenesis, three-dimensional ligand-binding domain modelling.

INTRODUCTION

Nuclear receptors constitute a superfamily of ligand-dependent transcription factors that regulate a wide array of physiological and vital processes in the metazoa, including growth, development, sexual differentiation and metabolism [1]. The activity of these proteins is regulated by small lipophilic molecules such as steroid hormones, retinoids, thyroid hormone or vitamin D, but the superfamily also includes a majority of members, called orphan nuclear receptors, for which no ligand has been found [2]. Natural ligands have recently been identified for several nuclear orphan receptors, as is the case for ROR α (retinoic acid-related orphan receptor α) [3], and it seems likely that the orphan group of nuclear receptors comprises both molecules with an as-yet-unidentified ligand, but also proteins that can function in a ligand-independent manner [4]. This has led to speculation about the origin of ligand binding within the nuclear receptor superfamily. The variety of ligands bound by members of the same subfamily, and the binding of chemically similar ligands by members of different receptor subfamilies, gave rise to the hypothesis that ligand binding was not ancestral, but rather was acquired during the evolution of nuclear receptors [5]. Whether or not they are orphans, nuclear receptors share a modular structure composed of six distinct regions designated A to F. No significant conservation between different nuclear receptors has been found for the A/B, D

and F domains. In many cases, the A/B region contains a ligand-independent transactivation function called AF-1 (activation function-1). The D domain serves as a flexible hinge between the C and E domains, and no particular functional conservation has been found for the F domain. The most conserved region among all the nuclear receptors is the C domain, or DBD (DNA-binding domain), which allows the binding of the molecule as a monomer or homo- or hetero-dimer to its DNA-response element through two zinc-finger modules. Most nuclear receptors act as dimers, either homodimers as in the case of steroid receptors, or as heterodimers with a member of the RXR (retinoid-X-receptor) family as a partner [1]. In addition, a few members of the superfamily bind DNA as monomer [SF-1 (steroidogenic factor-1), ROR and oestrogen-related receptor, for example]. The E domain is conserved to a lesser extent, and is the most complex and functionally important domain as it contains the LBP (ligand-binding pocket), the main dimerization interface and the ligand-dependent activation function (AF-2). Crystallographic studies of various nuclear receptor E domains in the presence of agonists or antagonists have solved the molecular mechanism of the ligand-dependent regulation of nuclear receptors [6,7]. These structural data revealed that the LBD (ligand binding domain) comprises 12 α -helices folded into a canonical three-layer 'helical sandwich', among which helix 12 contains the ligand-dependent activation function (AF-2; [8]). Unliganded or antagonist-bound

Abbreviations used: AF(1/2), activation function(-1/-2); DBD, DNA-binding domain; D⁵⁸⁵KF⁵⁸⁷/AAA, etc., a triple mutant bearing replacements of Asp⁵⁸⁵, Lys⁵⁸⁶ and Phe⁵⁸⁷ with alanine, etc; DS, dynamic stabilization; EcR, ecdysteroid receptor; ER, oestrogen receptor; GST, glutathione S-transferase; LBD, ligand-binding domain; LBP, ligand-binding pocket; LRH-1, liver receptor homologue-1; MAPK, mitogen-activated protein kinase; RAR, retinoid acid receptor; RID, receptor interacting domain; ROR α , retinoic acid-related orphan receptor α ; RXR, retinoid-X-receptor; SF-1, steroidogenic factor-1; SRC-1, steroid receptor coactivator-1; (Sm)Ftz-F1, (*Schistosoma mansoni*) Fushi Tarazu-Factor 1; 3D, three-dimensional; TR, thyroid hormone receptor.

¹ To whom correspondence should be addressed (email raymond.pierce@pasteur-lille.fr).

receptors are generally transcriptionally inactive, and interact with co-repressors such as SMRT (silencing mediator for retinoid and thyroid hormone receptor) and N-CoR (nuclear co-repressor) [9]. Ligand binding induces conformational changes and provokes the displacement of helix 12 that folds tightly against the LBD, creating a new surface where co-activators, such as SRC-1 (steroid receptor coactivator-1), can bind [7].

The Ftz-F1 (Fushi Tarazu-Factor 1) subfamily (nuclear receptor 5 subfamily under the unified nuclear receptor nomenclature [10]) contains orphan receptors that bind to their response element as monomers. The first member of the subfamily, Ftz-F1 α , was isolated from *Drosophila melanogaster* and identified both as a transcriptional regulator [11] and a cofactor [12] of the pair-rule class segmentation gene *fushi tarazu*. In the nematode *Caenorhabditis elegans*, nhr-25, the homologue of Ftz-F1, is required for epidermal and somatic gonad development and also participates in the regulation of moulting [13]. In mammals, an Ftz-F1 homologue was first identified as a steroidogenic factor [Ad4BP (adrenal 4 binding protein)/SF-1] present in the adrenal gland, and was able to bind to the proximal promoter region of cytochrome P450 steroid hydroxylase gene [14]. Further studies demonstrated that SF-1 is a key regulator of the endocrine function of the hypothalamus–pituitary–steroidogenic organ axis, as well as an essential factor in sex determination and gonadal differentiation (reviewed in [15]). A second subfamily of Ftz-F1, encoded by a separate gene named NR5A2, is represented by LRH-1 (liver receptor homologue-1) in the mouse [16], FTF (α -fetoprotein transcription factor) in the rat [17], PHR-1 (pancreas homologue receptor 1) in humans [18] and FF1 in zebrafish [19]. These receptors share identical structural and functional properties with members of the SF-1 subfamily, but are expressed in tissues that do not express SF-1, such as the liver, pancreas and intestine [17,18]. LRH-1 has been shown to be required for the regulation of critical genes involved in cholesterol metabolism and bile-acid synthesis [20].

The existence of an SF-1 or LRH-1 ligand has never been demonstrated, and the putative ligand 25-hydroxycholesterol was found to activate transcription only at very high concentrations, and also failed to increase transcription from a variety of SF-1-dependent promoters [21]. It is thought that post-translational modifications, notably the phosphorylation of the hinge at Ser-203 of SF-1, might be the most important mechanism controlling the activity of this orphan receptor [22]. Recently, the crystal structure of the LRH-1 LBD showed that this domain can adopt an active conformation without a ligand in the LBP and provided evidence for a ligand-independent transcription activity of the members of the SF-1/LRH-1 subfamily [4].

Ftz-F1 from the parasitic platyhelminth *Schistosoma mansoni* (SmFtz-F1) is the only member of the Ftz-F1 family of nuclear receptors characterized in a lophotrochozoan, and we demonstrated previously [23] that it possessed some conserved functional properties, notably in terms of DNA binding and the transactivational activity of the full-length protein. Despite this, structural differences, including the presence of a very large (104-amino-acid) F-domain, a poorly conserved E-domain and the absence of the consensus MAPK (mitogen-activated protein kinase) phosphorylation site [22], suggested that major functional differences also exist, especially when compared with mammalian SF-1. In the present study, we have characterized further the unique properties of the schistosome receptor by a systematic analysis of the functional domains of SmFtz-F1. Results obtained show that the D and F domains are essential for the AF-2-dependent activity of the receptor and argue for an important conformational role of these domains. To study the function of the LBD we have used both deletion analysis and site-directed muta-

genesis, but in contrast with previous studies, we have used 3D (three-dimensional) models of the LBD in order to determine those residues that are likely to be involved in receptor function rather than in the maintenance of the 3D structure of the domain. Combining this strategy with the LBD dynamic functional assay [24] allowed us to confirm the important role of helix 1 in maintaining the structural conformation of the LBD, and suggested that its autonomous transactivation activity, also observed in SF-1, is fortuitous. This strategy also allowed us to show that, unlike *Danio rerio* Ftz-F1 α , the SmFtz-F1 I-box is not implicated in SRC-1 recruitment. Finally, mutations of two arginine residues, whose side chains point towards the LBP, strongly decrease the AF-2-dependent activity of the receptor, possibly suggesting a ligand-dependent activity of SmFtz-F1. These unique and original functional properties of a schistosome nuclear receptor suggest that these molecules have evolved to acquire different and specific functions within different branches of the metazoa.

EXPERIMENTAL

Plasmid constructions, fusion proteins and deletions

PCR was used to create constructs expressing fusion proteins between the DBD of Gal4 (amino acids 1–147) [25] and SmFtz-F1 [23]. Platinum[®] Taq DNA polymerase High Fidelity (Invitrogen[™]) was used for all the PCRs, according to the manufacturer's instructions. The Gal4 DBD was cloned into the *EcoRI* site of pTL1 (a modified version of pSG5; Stratagene) to produce the pTL1Gal4 vector. All the constructs containing the SmFtz-F1 domains were generated by PCR using the pTL1FF1 expression vector [23] encoding the full-length SmFtz-F1 cDNA as template. All the primers used in the present study are listed in Supplementary Table 1 (<http://www.biochemJ.org/bj3810337add.htm>).

The fragment corresponding to the D, E and F domains of SmFtz-F1 was amplified with 5'-FF1 and 3'-FF1. Primers used to obtain deletions in the F domain were 5'FF1, 3' Δ F, 3' Δ F1 and 3' Δ F2. Primers used for deletion in the D domain were 5' Δ D1, 5' Δ D2, 5' Δ D3 and 3' FF1. All the PCR products were cloned into pCR4 (TOPO TA Cloning; Invitrogen) for sequence verification. Positive clones were digested with *HindIII* and *NotI*, and inserts were cloned in-frame with the Gal4 DBD into the pTL1Gal4 vector to generate constructs Gal-SmFF1, Gal- Δ F, Gal- Δ F1, Gal- Δ F2, Gal- Δ D1, Gal- Δ D2 and Gal- Δ D3.

Plasmids (Gal- Δ E1, Gal- Δ E2, Gal- Δ E3, Gal- Δ I box and Gal- Δ E4) containing internal deletions of segments encoding amino acids 383–593, 447–593, 514–593, 557–593 and 584–593 respectively were generated as follows: the SmFtz-F1 C-terminal fragment containing the AF-2-AD and F domains was digested with *EcoRI* and *NotI* from the pTL1FF1 vector and cloned into pBSK+. The SmFtz-F1 fragments from residue 249 to residues 383, 447, 514, 557 and 584 were obtained by PCR using 5'FF1 (forward primer) and 3' Δ E1, 3' Δ E2, 3' Δ E3, 3' Δ I box and 3' Δ E4 (reverse primers). Each fragment was cloned into pCR4 for sequencing. Positive clones were digested with *HindIII* and *EcoRI* and cloned in pBSK+ containing the *EcoRI/NotI* fragment. The resulting inserts were subcloned as *HindIII/NotI* fragments into the pTL1Gal4 vector. Constructs Gal- Δ E5 to Gal- Δ E9 were obtained from the same PCR products, but were first cloned into the *HindIII/EcoRI* sites of a wild-type pBSK+ vector (giving Δ E5pBSK+ to Δ E9pBSK+ vector) and then digested with *HindIII* and *PstI* and subcloned into the pTL1Gal4 vector. The fragment corresponding to amino acids 514–732 was amplified with primers 514–732 5' and 3'FF1. Following sequence verification, a positive clone was digested with *SmaI/NotI* and ligated into the Δ E5pBSK+ and Δ E6pBSK+ vectors. After digestion

with *HindIII/NotI*, the resulting fragments were subcloned into the pTL1Gal4 vector, giving constructs Gal- Δ E10 and Gal- Δ E12 respectively. Construct Gal- Δ E11 was obtained using the same strategy with a PCR product corresponding to amino acids 447–732 first cloned into the Δ E5pBSK+ vector before subcloning into pTL1Gal4. cDNA fragments encoding amino acids 383–447 and 557–732 were amplified using primer pairs 5'FF1/383–447 3' and 557–732 5'/3'FF1. Each PCR product was cloned into pCR4, sequenced, and positive clones were subcloned into pTL1Gal4 using the restriction sites *HindIII/NotI* introduced in the primers.

A human SRC-1 expression vector (pCR3.1-hSRC1a) was kindly provided by B. W. O'Malley (Baylor College of Medicine, Houston, TX, U.S.A.). To generate the GST (glutathione-S-transferase)–SRC-1 fusion, we amplified the receptor-interacting region of SRC-1 (from residue 1243 to the C-terminus) [26] with primer SRC1-ID 5' and primer SRC1-ID 3' using pCR3.1-hSRC1a as template. After sequence verification, the PCR product was cloned in-frame with GST into the *BamHI/NotI* cloning sites of the pGEX-4T3 vector (Amersham Biosciences).

The expression vectors encoding the fusion protein between VP16 and human SRC1 ID [from residue 1243 to the C-terminus; pAASV-VP16hSRC1(.8)] and the VP16 protein alone (pAASV-VP16) were kindly provided by Dr J. L. Jameson (Northwestern University Medical School, Chicago, IL, U.S.A.).

VP16 fusion proteins for the LBD DS (dynamic stabilization) assay were generated by PCR using pTL1FF1 as template with different oligonucleotide combinations. The VP16 410–632 construct was obtained using primers DS5'1 and DS3', the VP16 410–731 construct was obtained using DS5'1 and DS3' + F, the VP16 438–632 construct was obtained using DS5'2 and DS3' and the VP16 438–731 construct was obtained using DS5'2 and DS3' + F. After sequence verification, inserts of positive clones were cloned in-frame with the VP16 activation domain into the *BamHI/HindIII* restriction sites of the pVP16 vector (ClonTech).

Site-directed mutagenesis

For mutagenesis of the SmFtz-F1 AF-2-AD and F domain, we used a PCR-based site-directed mutagenesis strategy, using pTL1FF1 as a template and Platinum[®] *Taq* DNA polymerase High Fidelity[™] (Invitrogen). A couple of complementary primers was designed for each mutated construct. Primers were AF-2mutFw/AF-2mutRev for AF-2-AD and FmutFw/FmutRev for the F domain. For both mutageneses, the first step was to obtain, by PCR, two fragments between 5'FF1/mut Rev primer and mut Fw primer/3'FF1. Fragments were fused using 8 μ l of each PCR product and an amplification procedure, consisting of 3 min at 94 °C followed by 5 cycles at 94 °C (5 s), 60 °C (25 min) and 68 °C (5 min). An aliquot (1 μ l) of a 1:100 dilution of this fusion PCR was re-amplified using primers corresponding to the extremities of the fused fragment (5'FF1 and 3'FF1). Following cloning into pCR4 and verification by sequencing, mutated fragments were cloned in the pTL1Gal4 vector after digestion by *HindIII/NotI*.

The SmFF1 I-box mutants {[D⁵⁸⁵KF⁵⁸⁷/AAA (a triple mutant bearing replacements of Asp⁵⁸⁵, Lys⁵⁸⁶ and Phe⁵⁸⁷ with alanine)], K⁵⁸⁹LI⁵⁹¹/AAA, D585A, K586A, F587A, N588A, N592A and E596A}, the LBP mutants (R448A, R448T, R489A, A445V, A445F, C486V and A604V), the H1 mutants and the Gal-AF-2mut2 construct were generated using the QuikChange[®] Site-directed Mutagenesis Kit (Stratagene), according to the manufacturer's instructions. Briefly, Gal4SmFF1 was used as a template (except for two mutants; see below) and the oligonucleotides used were DKF/AAA5' and DKF/AAA3' for the DKF mutant; KLI/

AAA5' and KLI/AAA3' for the KLI mutant; D/A5' and D/A3' for the D585A mutant; K/A5' and K/A3' for the K586A mutant; F/A5' and F/A3' for the F587A mutant; N1/A5' and N1/A3' for the N588A mutant; N2/A5' and N2/A3' for the N592A mutant; E/A5' and E/A3' for the E596A mutant; R448/A5' and R448/A3' for the R448A mutant; R448/T5' and R448/T3' for the R448T mutant; R489/A5' and R489/A3' for the R489A mutant; A445/V5' and A445/V3' for the A445V mutant; A445/F5' and A445/F3' for the A445F mutant; C486/V5' and C486/V3' for the C486V mutant; A604/V5' and A604/V3' for the A604V mutant; H1mut5' and H1mut3' for the Gal-SmFF1mutH1 and Gal- Δ E6mutH1 constructs; and AF-2-E/A5' and AF-2-E/A3' for the Gal-AF-2mut2 construct. Oligonucleotides containing the desired mutations were annealed to the denatured template and extended during temperature cycling by *PfuTurbo* DNA polymerase. The parental DNA template was digested with *DpnI*, and then the mutated plasmids were transformed into XL1-Blue 'super-competent' cells. Following sequence verification, positive clones were used directly in transfection assays. The double-arginine mutant was generated using the same method, with Gal-R448A cDNA as template and oligonucleotides R489/A5' and R489/A3'. Gal- Δ E6 cDNA was used as a template for the Gal- Δ E6mutH1 construction. For the LBD DS assay, a stop codon was introduced using the same strategy. The Gal 249–437 construct was obtained with oligonucleotides GalH1H2 5' and GalH1H2 3' using Gal- Δ E6 cDNA as template. The same primers were used with Gal-H123 cDNA to obtain the Gal 383–437 construct. Oligonucleotides GalH1 5' and GalH1 3' were used with Gal- Δ E6 cDNA and Gal-H123 cDNA as templates, to obtain the Gal 249–409 and Gal 383–409 constructs respectively.

Cell culture and transfection

CV-1 cells were maintained in Dulbecco's modified Eagle's medium supplemented with 10 % (v/v) fetal bovine serum and 50 μ g/ml gentamicin at 37 °C and 5 % CO₂. Cells were plated 1 day before transfection in six-well plates at a density of 200 000 cells/well. Transfections were performed with linear PEI ExGen 500 (Euromedex, Souffelweyersheim, France), under the conditions recommended by the supplier. Cells were lysed 48 h after transfection, and were assayed using a luciferase assay kit (Promega) in a Wallac Victor² 1420 multilabel counter (Perkin-Elmer). For all experiments, protein content was used to normalize luciferase results. Transcriptional activities and interaction properties of the Gal4 fusion proteins were assessed on a luciferase reporter gene driven by five copies of the Gal4-binding site upstream of a basic promoter element (TATA box; pFR Luc plasmid; Stratagene). Each experiment represents at least three sets of independent triplicates.

Expression of GST fusion protein and GST pull-down assay

An empty GST construct (pGEX 4T3) and GST–SRC1 were used to transform the BL21 strain of *Escherichia coli* and proteins were purified as described previously [27]. Briefly, protein expression was induced with 1 mM IPTG (isopropyl β -D-thiogalactoside) at 30 °C for 3 h. Cells (corresponding to a culture of 50 ml) were harvested, resuspended in 2.5 ml of cold PBS containing protease inhibitors (1 mM PMSF and protease-inhibitor cocktail; Sigma) and then sonicated for 2 $\times$ 1 min. Triton X-100 was added to a final concentration of 1 %, and lysates were incubated for 30 min on ice with frequent shaking. The suspensions were clarified by centrifugation, and the supernatants were immediately placed in contact with 60 μ l of a 50 % suspension of glutathione–agarose beads (Sigma) in PBS. Incubation was performed

overnight under rotation at 4 °C. Protein bound to the agarose beads was washed twice with 300 μ l of cold PBS, and twice with 300 μ l of binding buffer [20 mM Tris/HCl (pH 8.0)/100 mM NaCl/1 mM EDTA/0.02 % (w/v) BSA/0.02 % (v/v) Tween 20/1 mM PMSF/protease inhibitor cocktail (Sigma)]. [35 S]-Methionine-labelled proteins (5 μ l) translated *in vitro* using the Quick Coupled Transcription/Translation System (Promega) were added to the washed beads in a final volume of 200 μ l of binding buffer, and incubated for 2 h at 4 °C under rotation. Beads were collected by centrifugation, washed three times with 300 μ l of binding buffer, resuspended in SDS/PAGE sample buffer and boiled for 5 min before loading on a 10 % denaturing polyacrylamide gel. The gels were fixed (10 % acetic acid/10 % ethanol), dried and exposed to β Max film (Kodak) overnight.

LBD modelling

Models of the *S. mansoni* FTZ-F1 (SmFtz-F1) LBD were generated by homology modelling using the human RAR γ (retinoid acid receptor γ) (2LBD; [6]), *Heliothis* EcR (ecdysteroid receptor) (1R1K; [28]) and mouse LRH-1 (1PK5; [4]) LBD structures as references with the Modeller software, using default parameters [29]. The quality of the homology models was evaluated by Procheck [30] and ProsaII [31]. The stereochemical parameters calculated by Procheck show that more than 97 % of the residues in all three models have the Φ/ψ dihedral angles in the most favoured region of the Ramachandran plot. In addition, ProsaII gives combined Z-scores (pair and surface potentials) of -6.6, -7.26 and -7.7 for the models respectively based on RAR, EcR and LRH-1. These values, which are close to the range observed for experimental nuclear receptor structures (located between -7 and -10), indicate a rather good model quality.

RESULTS

SmFtz-F1 interacts with human SRC-1

To study the transcription activation domains of SmFtz-F1 independently of its DNA binding and nuclear localization properties, we used a series of fusion proteins between the DBD of the yeast transcription factor Gal4 and different deletion mutants of the receptor. Each construct was co-transfected into CV-1 cells with the reporter plasmid pFR-Luc (Stratagene), which contains five Gal4 binding sites upstream of a minimal thymidine kinase promoter and the firefly luciferase gene. The equivalent expression of each construct in transfected cells was monitored by Western blotting using an anti-Gal4DBD antibody (results not shown). The Gal-SmFF1 construct, containing the D, E and F domains of SmFtz-F1, led to an approx. 10-fold induction of the reporter gene compared with the activation obtained with pTL1-Gal4, which contains only the Gal4 DBD (amino acids 1–147; Figure 1A). This finding showed that, even though CV-1 cells represent a heterologous system for the study of a schistosome nuclear receptor, significant transactivation of the reporter gene was obtained, indicating that at least some of the mammalian co-activators interact with SmFtz-F1, most probably through its conserved AF-2 domain. To test this hypothesis and to characterize further SmFtz-F1, we asked whether or not SRC-1 could interact with the receptor. SRC-1 has been identified as a ligand-dependent co-activator of a broad range of nuclear receptors, including steroid and non-steroid receptors [32], but also orphan receptors such as SF-1 [33]. Because no nuclear receptor cofactors have been cloned from *S. mansoni* so far, we have used the RID (receptor-interacting domain) of the human SRC-1 protein fused to the VP16 activation domain (kindly provided by Dr L. Jameson) in a mammalian

Figure 1 Transcriptional activity of the Gal4-SmFF1 construct and interaction properties with the SRC1 RID

(A) The transcriptional activity of SmFtz-F1 domains D, E, F fused to the Gal4 DBD (amino acids 1–147) was examined in CV-1 cells by co-transfecting the cells with 200 ng of pFR-LUC reporter plasmid and 800 ng of Gal4 expression vector. Results are expressed as relative normalized luciferase activity (means $\pm$ S.D.). (B) Interaction between SmFtz-F1 and SRC1 was assessed in CV-1 cells using a mammalian two-hybrid assay. Cells were transfected with 200 ng of pFR-LUC reporter plasmid, 200 ng of Gal4 expression vector and increasing amounts of VP16-SRC1RID expression vector as indicated. Results are expressed as the fold activity relative to the basal reporter activity obtained with co-transfection with plasmid encoding GAL4 and VP16 alone. (C) SmFtz-F1 interacts with SRC1 RID *in vitro*. A GST pull-down assay was used to confirm interaction between SmFtz-F1 and SRC1 RID *in vitro*. Glutathione beads, bound with GST alone or the GST-SRC1RID construct, were incubated with 5 μ l of *in vitro*-translated [35 S]methionine-labelled Gal-SmFF1 construct for 2 h at 4 °C. After washing extensively, the proteins bound on beads were analysed by SDS/PAGE and visualized by autoradiography.

two-hybrid system. For this purpose, using the same reporter vector (pFR-Luc), either the VP16 or VP16-SRC-1 expression vector were co-transfected with Gal4 alone or the Gal4-SmFF1 construct. With VP16 alone, Gal-SmFF1 transcriptional activation was approx. 10-fold higher than background. Co-transfection with SRC1-VP16 strongly enhanced the transactivation activity of Gal-SmFF1 in a dose-dependent manner (Figure 1B), reflecting the recruitment of the VP16 activation domain by an interaction between SmFtz-F1 and the human SRC-1 RID. Remarkably, the potentiation of the Gal-SmFF1 transactivation activity by VP16-SRC1 is very strong (approx. 60 fold with 800 ng of the VP16-SRC1 construct) and appears greater than the potentiation obtained for other members of the subfamily in the same system [33]. This result was confirmed *in vitro* by a GST pull-down assay using hSRC-1 RID fused to GST. As shown in Figure 1(C), Gal-SmFF1 interacts specifically with hSRC-1 *in vitro* and no interaction was detected with GST alone [the presence of a second, lower-molecular-mass band is uniquely due to the fortuitous use of ATG within the Gal4DBD-coding region (not shown) and both

bands represent the recombinant SmFtz-F1]. We had previously checked that the Gal4 DBD does not interact with SRC-1 RID *in vitro* (results not shown). Taken together, these results indicate that, when fused to the Gal4 DBD, the D, E and F domains of SmFtz-F1 are sufficient for transactivation and interaction with cofactors in CV-1 cells.

The AF-2-AD domain is the major activation function of SmFtz-F1

We next decided to assess the functional importance of the AF-2 domain of SmFtz-F1. This domain is very well conserved, and perfectly conforms to the consensus sequence established previously [34] (Figure 2A). Two different mutants were made: in Gal-AF-2mut1, the first two leucine residues (Leu-622 and Leu-623) of the AF-2 core motif were mutated to alanine, since such mutations were previously shown to strongly inhibit the AF-2 activity of mouse SF-1 [35]. In Gal-AF-2mut2, the conserved glutamate (Glu-625) was mutated to alanine (Figure 2A). This residue, together with a basic residue (here, Arg-457), has been shown to form a clamp with the LXXLL (Leu-Xaa-Xaa-Leu-Leu) motif of the co-activator in the groove formed by helices 3, 4, 5 and 12 [7]. The activities of these mutants were examined in CV-1 cells using pFR-Luc as a reporter plasmid. As shown in Figure 2(B), luciferase activity obtained with both mutants was reduced to the background level, indicating a complete loss of transcriptional activation by SmFtz-F1. This result suggests that AF-2-AD is critical for the transactivation function of SmFtz-F1. This was next confirmed by examining the interaction between both mutants and the SRC-1 RID in our mammalian two-hybrid system (Figure 2C). Indeed, as described previously, Gal-SmFF1 strongly interacts with SRC1-RID, whereas co-transfection of each mutant either with VP16 alone or the VP16-SRC1 construct resulted in a background level of luciferase activity, indicating that there was no interaction between the Gal-AF-2 mutants and the SRC-1 RID. The GST pull-down assay also demonstrated that no interaction exists between AF-2 mutants and SRC1 *in vitro* (Figure 2D). These results confirm our hypothesis and indicate that the conserved AF-2 domain of SmFtz-F1 is essential for receptor activity and that, as for members of the liganded steroid receptor family, the effect of SRC1 on SmFtz-F1 is mediated through the AF-2 domain.

The F domain is important for transactivation

Not all nuclear receptors contain an F domain. Where its function is known, this region has been shown to display very distinct and specific roles for each receptor. SmFtz-F1 has the longest F domain (104 amino acids) among all the members of the Ftz-F1 family, in which this domain is often lacking. Moreover, when it is present, this domain is not conserved among the Ftz-F1 proteins from different species. It was intriguing that SmFtz-F1 should possess such a long F domain, and we therefore assessed its importance in the transcriptional activity of the receptor. For this purpose, we tested a series of C-terminal deletions fused to the Gal4 DBD for their transactivation activity in CV-1 cells (Figure 3A). Deletion of the F domain results in a 50% reduction in transcriptional activity of Gal-SmFF1 (construct Gal-ΔF), indicating that this region is important for the full activity of SmFF1. A motif that may be involved in this activity is an LXXLL-related sequence in the F domain between residues 690–694 (V⁶⁹⁰TDLL⁶⁹⁴). It was recently reported that LXXLL-related sequences present in the N-terminal domain of the androgen receptor interact with the C-terminal domain, notably with the AF-2-AD [36]. It thus seemed possible that the loss of trans-

Figure 2 The AF-2-AD is crucial for the transcriptional activity of SmFtz-F1

(A) Residues mutated in the two mutAF-2 Gal4 fusion proteins. The consensus sequence of the AF-2-AD core is represented ('Φ' denotes a hydrophobic residue, 'E' is glutamic acid and 'X' any amino acid). The AF-2-AD sequence of SmFtz-F1 conforms exactly to this consensus (amino acid numbers are indicated). The first two leucine residues were mutated to alanine in the Gal-AF-2mut1 construct and the conserved glutamic acid was changed to alanine in the Gal-AF-2mut2 construct. (B) The Gal-mutAF-2 constructs are transcriptionally inactive. CV-1 cells were co-transfected with 200 ng of pFR-Luc reporter vector and 800 ng of Gal4 expression vector. The results are expressed as the percentage of normalized luciferase activity with the activity of the Gal-SmFF1 construct taken as 100%. (C) AF-2-AD mutations abolish interaction with SRC1-RID in CV-1 cells. Cells were transfected with 200 ng of pFR-LUC reporter plasmid, 200 ng of Gal4 expression vector and 800 ng of VP16-SRC1 RID expression vector. Results are expressed as the fold activity relative to the basal reporter activity obtained in a co-transfection with plasmid encoding GAL4 and VP16 alone. (D) AF-2-AD mutations abolish interaction with SRC1-RID *in vitro*. Aliquots (5 μl) of *in vitro*-translated [³⁵S]methionine-labelled Gal-mutAF-2 constructs were used in a GST-pull down assay as described in the Experimental section.

criptional activity with the construct Gal-ΔF may have been due to improper AF-2 orientation, in turn caused by the lack of interaction between the F domain and AF-2 through the VTDLL motif. To test this hypothesis, we replaced the two leucine residues of the VTDLL sequence with alanine (construct Gal-mutF), since

Figure 3 SmFtz-F1 requires its F domain for full activity

(A) Transcriptional activity of constructs deleted or mutated in the F domain fused to Gal4 DBD compared with the Gal-SmFF1 fusion protein. A schematic representation of each construct is depicted, and amino acid positions are indicated. CV-1 cells were transfected with 200 ng of pFR-Luc reporter vector and 800 ng of Gal4 expression vector. The results are expressed as the percentage of normalized luciferase activity, with the activity of the Gal-SmFF1 construct taken as 100%. (B) The F domain influences SRC1-RID interaction with the AF-2-AD. CV-1 cells were transfected with 200 ng of pFR-LUC reporter plasmid, 200 ng of Gal4 expression vector and 800 ng of VP16-SRC1RID expression vector. Results are expressed as the fold activity relative to the basal reporter activity obtained after co-transfection with the plasmids encoding GAL4 and VP16 alone.

such mutations were shown to abolish interaction of the LXXLL motif with AF-2 [37], and we also attempted to sub-localize the transactivation potential of the F domain further by testing deleted constructs Gal-ΔF1 and Gal-ΔF2. The results in Figure 3(A) show that none of the three constructs had a significantly impaired transcriptional activity compared with Gal-SmFF1, indicating that the V⁶⁹⁰TDLL⁶⁹⁴ motif seems not to play a particular role, and if residues of the F domain are essential for transactivational activity, they are located in the first 50 amino acids of this domain.

To study further the influence of the F domain on AF-2 activity of SmFtz-F1, we used the mammalian two-hybrid system with the SRC-1 RID fused to the VP16 activation domain (Figure 3B). The interaction between SRC1 and the receptor lacking its F domain (Gal-ΔF) was strongly impaired but still present (200-fold compared with 1400-fold induction), showing that the F domain influences the AF-2 activity of SmFtz-F1. The two C-terminal-deleted proteins showed a reduced interaction, intermediate between the wild-type and the Gal-ΔF activities. Taken together, our results indicate that the F domain of SmFtz-F1 is essential for the full transcriptional capacity of the receptor, and strongly contributes to the AF-2-dependent activity of the protein.

The hinge region is essential for SmFtz-F1 activity

No significant peptide sequence conservation among different nuclear receptor families was found for the D domain, and its role remains elusive. It was first considered as a flexible hinge between the DBD and the LBD, allowing some receptors to bind as dimers

to both direct and inverted repeat response elements. Recent studies have suggested that the D region is more than simply a flexible joint, and that the integrity of the whole domain is necessary to maintain an active conformation of the receptor [38,39]. In the case of SF-1, the phosphorylation of a serine residue present in the hinge domain increased the stability of the LBD [22]. To analyse whether the hinge domain of SmFtz-F1 is involved in the transcriptional activity of the receptor, we tested a series of N-terminal deletions of the D domain fused to the Gal4 DBD for their potential to transactivate the pFR-Luc reporter vector (Figure 4). Co-transfection with the mutant Gal-ΔD1, which lacks the entire D domain, resulted in the loss of transactivation compared with Gal-SmFF1 (7% of Gal-SmFF1 activity). A background level of luciferase activity was maintained upon transfection with the construction Gal-ΔD2, which lacks residues 249–342 (8% of Gal-SmFF1 activity), and a low level of transactivation was retrieved with Gal-ΔD3 which lacks residues 249–295 (22% of Gal-SmFF1 activity). These results seem to indicate that the whole hinge region is essential for the transcriptional activity of SmFtz-F1, probably by promoting the active conformation of the receptor. This result was confirmed by testing the three truncated proteins in the mammalian two-hybrid system with the SRC-1 RID (Figure 5A). Whereas Gal-ΔD3 interacted quite well with SRC-1 RID, Gal-ΔD2 and Gal-ΔD1 interacted only weakly, probably indicating an improper folding of helix 12 and of the entire LBD. This two-hybrid assay argues for an important conformational role for the D domain, which seems to be essential for the maintenance of the integrity of the functional conformation of SmFtz-F1, but also indicates that amino acid residues 249–295 could be involved in another mechanism. Indeed, whereas 70% of the wild-type SRC-1 interaction is observed with the Gal-ΔD3 construct (Figure 5A), this deleted protein only showed a low level of transactivation (22% of Gal-SmFF1 activity, as shown in Figure 4), suggesting a possible role of residues 249–295 in transcriptional regulation.

Regulatory role of SmFtz-F1 LBD sequences

The E domain of nuclear receptors is very complex, since it contains the LBD, harbours a strong dimerization interface and the ligand-dependent transactivation function AF-2. To investigate the importance of this domain for a monomeric orphan nuclear receptor and to study the relationship between the D and E domains, we fused a series of N-terminal and C-terminal deletions of the LBD to the Gal4DBD, and tested their transcriptional activities (Figure 4). Constructs Gal-ΔE1 to Gal-ΔE4 still possess at least the putative helix 12 (AF-2-AD) of the LBD and also the F domain. Surprisingly, co-transfection with mutants Gal-ΔE1 and Gal-ΔE2 resulted in 2.6- and 7-fold increases in the Gal-SmFF1 activity respectively. Adding the following 67 residues (amino acids 447–514) abolished this effect, and the luciferase activity stimulated by mutant Gal-ΔE3 was reduced to 50% of the Gal-SmFF1 activity. A low level of transactivation was maintained upon transfection with the deleted constructs Gal-ΔIbox and Gal-ΔE4. We next tested whether this effect was AF-2-dependent or -independent by using our mammalian two-hybrid assay. As shown in Figure 5(B), none of the mutants could interact with SRC1, suggesting that the AF-2 core was not in an active conformation in these five deleted proteins and that the effect observed with Gal-ΔE1 and Gal-ΔE2 was AF-2-independent. To confirm this hypothesis, we co-transfected in CV-1 cells along with pFR-Luc the same series of truncated constructs, but lacking the C-terminal part corresponding to helix 12 and the F domain (constructs Gal-ΔE5 to Gal-ΔE9; Figure 4). Remarkably, mutant Gal-ΔE5, which only contains the D domain,

Figure 4 Transcriptional activity of constructs deleted in the D and E domains of SmFtz-F1 fused to Gal4 DBD compared with the Gal-SmFF1 construct

A schematic representation of each construct is depicted and amino acid positions are indicated. CV-1 cells were transfected with 200 ng of pFR-Luc reporter vector and 800 ng of Gal4 expression vector. The results are expressed as the percentage of normalized luciferase activity, with the activity of the Gal-SmFF1 construct taken as 100%.

still showed significant transcriptional activity corresponding to approx. 50% of Gal-SmFF1 activity. The construct Gal-ΔE6, containing the D domain and amino acids 383–447 of the E domain, showed a remarkable 16-fold increase in activity compared with Gal-SmFF1. As shown previously, this effect was strongly reduced in mutants containing additional residues corresponding to amino acids 447–514 (construct Gal-ΔE7), and the transactivational activity was not retrieved with constructs Gal-ΔE8 and Gal-ΔE9. These results clearly suggest that residues 383–447 harbour a strong autonomous transactivation domain that is inhibited by residues situated between positions 447 and 514. To characterize further this phenomenon, we tested N-terminal deletions of the LBD. To test the hypothesis that the inhibition observed with residues 447–514 was due to the binding of co-repressor molecules, we used mutants Gal-ΔE10, Gal-ΔE11 and Gal-ΔE12 (Figure 4). Construct Gal-ΔE12, containing an internal deletion corresponding to residues 447–514, which were apparently responsible for the negative regulation, nevertheless showed a reduced transcriptional activity compared with Gal-SmFF1. Moreover, the transactivational activity of this mutant is comparable with those obtained with constructs Gal-ΔE7 and Gal-ΔE3, which contain residues 447–514. Finally, the transcriptional activity did not change significantly when this latter region was added N-terminally in construct Gal-ΔE11 (compare Gal-ΔE10 and Gal-ΔE11). This result therefore clearly suggests that the region encompassing residues 447–514 does not contain an eventual repressor binding site that could explain the inhibition observed between mutants Gal-ΔE2 and Gal-ΔE3 or mutants Gal-ΔE6 and Gal-ΔE7. Moreover, none of the deleted proteins Gal-ΔE10, Gal-ΔE11 or Gal-ΔE12 were capable of interacting with SRC1 in the mammalian two-hybrid assay (Figure 5B), indicating that helix 12 is probably not correctly folded against the LBD in these mutants. All interaction experiments were confirmed *in vitro* by the GST-pull down assay (results not shown). Finally, we tested the activity of residues 383–447 without the D domain and found

that this construct, called Gal-ΔE13, showed a reduced luciferase activity compared with Gal-ΔE6, but which is still 3-fold higher than that obtained for Gal-SmFF1. Recently, Desclozeaux et al. [22] have demonstrated the importance of helix 1 of the LBD and the hinge region for stabilization of the LBD of SF-1. They also suggested the existence of an autonomous activation function in helix 1 of the LBD. In an identical manner, our results have identified an autonomous activation function in the N-terminal region of the SmFtz-F1 LBD containing residues 383–447. This function can act independently, but the hinge region strongly potentiates its effect. The activity of this domain is inhibited by the proceeding residues of the molecule, which do not encompass a repressor binding site. We were intrigued by these results, which raise the question of how the autonomous activation function acts in the context of the full-length protein, since it seems to be inactive when proceeding residues are present. In the light of our results, we can hypothesize that the transactivation activity of the Gal-ΔE6 construct, which encompasses residues 249–447, is likely to be fortuitous. Indeed, the 383–447 region could possess residues that serve to recruit powerful endogenous transcriptional co-activators. In the construct Gal-ΔE7, the additional residues could probably mask some amino acids, preventing binding of opportunistic endogenous co-activators. However, in order to improve our understanding of the suggested mechanisms, we needed to better appreciate the overall structural organization of the LBD of SmFtz-F1.

3D models of the LBD

We decided to model the 3D structure of this domain in order to rationalize the functional study of the SmFtz-F1 LBD, and to clarify our results and hypotheses concerning the LBD of SmFtz-F1. The purpose was to determine whether the overall structure of the LBD is conserved and to explain the results obtained by co-transfection experiments with deletion mutants. Three different

Figure 5 Subregions of both the D and E domains are required for full interaction of SmFtz-F1 with SRC1-RID

(A) Interaction between the three D domain-deleted constructs and SRC1-RID. CV-1 cells were transfected with 200 ng of the pFR-LUC reporter plasmid, 200 ng of Gal4 expression vector and 800 ng of VP16-SRC1RID expression vector. Results are expressed as fold activity relative to the basal reporter activity obtained by co-transfection with plasmids encoding GAL4 and VP16 alone. (B) Interaction between the E domain-deleted constructs and SRC1-RID. CV-1 cells were transfected with 200 ng of pFR-LUC reporter plasmid, 200 ng of Gal4 expression vector and 800 ng of VP16-SRC1RID expression vector. Results are expressed as shown in (A).

models were built: SmFtz-F1rar, SmFtz-F1ecr and SmFtz-F1lrh, based on the human RAR γ , *Heliothis* EcR and mouse LRH-1 crystal structures respectively. All three models (available on request from S. S. and J. M. W.) show a good overall conservation of the 12 α -helix LBD structure [8], except for helix 2. Indeed, a nearly complete lack of regulatory secondary structural elements was predicted for H2 in SmFtz-F1 (using the PHD program; [40]). Moreover, in the model SmFtz-F1lrh, the interaction of H2 with H3 is not maintained. Arg-352, for example, which is seen to be important for stabilization by H2, is not conserved in SmFtz-F1. Interestingly, SmFtz-F1 LBD exhibits a large and hydrophobic cavity [689, 690 and 777 Å³ (1 Å³ $\equiv$ 0.1 nm³) in the SmFtz-F1rar, SmFtz-F1ecr and SmFtz-F1lrh models respectively], which may correspond to an LBP.

Region 383–447 shows a fortuitous activation function

We next attempted to determine whether the activity displayed by construct Gal-ΔE6 (Figure 4) was fortuitous, or reflected the presence of a real autonomous activation function. Our hypothesis was that residues 383–447 contain motifs that might be responsible for the recruitment of powerful endogenous transcriptional co-activators. In the context of Gal-ΔE7 or the full-length protein, these motifs are masked, preventing cofactor binding. If this is the case, such motifs would not represent an autonomous transactivation function, since they would normally be buried inside the LBD, inaccessible to co-activators. The predicted 3D model of the LBD of SmFtz-F1 allowed us to determine the position of the putative helices, and residues 383–447 contain the entire helix 1 and the first half of helix 3. We noted that

Figure 6 Transcriptional activity of region 383–447 is in part carried by helix 1 and appears to be fortuitous

(A) The sequence corresponding to amino acids 384–415 of SmFtz-F1 is represented. Residues corresponding to the predicted helix 1 are boxed and predicted buried residues of helix 1 are underlined. The core consensus sequence of the AF-2-AD-related motif in helix 1 (' Φ ' denotes a hydrophobic residue, 'E' is glutamic acid and 'X' any amino acid). The mutated residues in helix 1 (in constructs Gal-SmFF1mutH1 and Gal- Δ E6mutH1) are represented. (B) Details of the interactions stabilizing the helix H1 in the SmFtz-F1 model. The contacts are due to residues in helices H3 and H8. Protein is depicted as a backbone carbon Ca trace. Atoms are represented in a standard colour scheme: carbon in grey, nitrogen in blue and oxygen in red. The side chains of residues participating in the interactions are labelled. Similarly to the RXR α and RAR γ structure [8], on one side, Asp-411 in H1 forms a hydrogen bond with Arg-451 in H3. On the other side, the H1 is stabilized by a series of contacts, involving a hydrogen bond between Glu-409 and Lys-543 and a van der Waals contact between Ile-406 and Ile-536. In the whole LBD structure, Asp-398, Leu-401 and Val-402 are hidden by H9. (C) Transcriptional activity of constructs carrying mutations D398A, L401G and V402G. CV-1 cells were transfected with 200 ng of pFR-Luc reporter vector and 800 ng of Gal4 expression vector. The results are expressed as the percentage of normalized luciferase activity, with the activity of the Gal-SmFF1 construct taken as 100 %.

two AF-2-AD-related motifs are present in the stretch of amino acids corresponding to helix 1 (Figure 6A). The first motif, IIRDIL, conforms exactly to the consensus sequence $\Phi\Phi XE\Phi\Phi$ (where ' Φ ' denotes a hydrophobic residue, 'E' is glutamic acid and 'X' any amino acid), except for the glutamic acid residue, which is substituted with aspartic acid. The second motif, LVEETI, also conforms to the consensus sequence except for the threonine residue. Moreover, in the whole LBD structure, seven residues of putative helix 1 are predicted to be buried (Figure 6B; underlined residues in Figure 6A), including Asp-398, Leu-401 and Val-402, which belong to the AF-2-AD-related motifs. Thus we decided to test our hypothesis by mutating these residues in the construct Gal- Δ E6, and also in Gal-SmFF1, giving the Gal- Δ E6mut H1 and Gal-SmFF1mutH1 constructs respectively. Aspartic acid was replaced by alanine and the leucine and valine residues were replaced by glycine residues (Figure 6A). Results are shown in Figure 6(C). The Gal- Δ E6 construct shows the same transcriptional capacity as described previously. As expected, the Gal- Δ E6mutH1 construct showed a transactivational activity that was reduced by approx. 60 % compared with the Gal- Δ E6 construct, indicating that the recruitment of co-activators is reduced by the mutation of the AF-2-related motifs. Mutations in helix H1 also impaired the transcriptional capacity of the Gal-SmFF1 construct, since the mutant Gal-SmFF1mutH1 showed a 70 % reduced luciferase activity compared with the wild-type Gal-SmFF1. This is probably due to the disruption of the interaction between helix 1 and the remainder of the LBD. Therefore our results clearly show that helix 1 harbours AF-2-AD-

related motifs that contain residues which participate in the powerful transcriptional activation mediated by the Gal- Δ E6 construct. Moreover, these residues are predicted to be buried inside the LBD, explaining why this activation is completely abolished when the following helices are added (in Gal- Δ E7, for example). The transactivation activity of the Gal-SmFF1mutH1 mutant clearly suggested that these H1 residues should be involved in intramolecular contacts essential to maintain the structural conformation of the LBD. Thus we may conclude from our results that there is no genuine autonomous transactivation function in the N-terminal part of the SmFtz-F1 LBD.

Helix 1 interacts with the remainder of the LBD

To confirm this result, we used the recently described LBD DS assay, which tests the interaction between helices 1–2 and the remainder of the LBD [24,41]. The predicted 3D model of the LBD of SmFtz-F1 indicates that helix 1 occurs between residues 397 and 409 and that there is no real helix 2, but that the region encompassing residues 409 and 437 appears to be unstructured. We therefore constructed four different Gal4 fusion proteins (Figure 7A) that would allow us to determine the role of helix 1 and the 409–437 region in the DS assay, and also to confirm the importance of the hinge region. Four VP16 fusion proteins were also made (Figure 7A), each of which contains residues 437–628 of the SmFtz-F1 LBD, which corresponds to the predicted helices 3–12. Two constructs possess the 409–437 unstructured region, and the influence of the F domain on the DS assay was also

Figure 7 SmFtz-F1 LBD DS assay: the region encompassing the hinge and helix H1 interacts with the remainder of the LBD

(A) Schematic representation of the SmFtz-F1 constructs used in the assay. (B) Mammalian two-hybrid assay. CV-1 cells were transfected with 200 ng of pFR-LUC reporter plasmid, 200 ng of Gal4 fusion expression vector and 800 ng of VP16 fusion expression vector. Results are expressed as the fold activity relative to the basal reporter activity obtained by co-transfection with plasmid encoding GAL4 and VP16 alone.

determined. The ability of these different proteins to interact was tested in the mammalian two-hybrid system. As shown in Figure 7(B), the induction of the reporter gene with construct Gal 249–437 alone (co-transfected only with VP16) was approx. 10-fold higher than with Gal4 alone, which is in good agreement

with our previous results. Indeed, this construct possesses helix H1 and this transcriptional activation reflected the recruitment of fortuitous endogenous co-activators by motifs contained in this helix. When the Gal 249–437 construct was co-transfected with the VP16 438–632 plasmid the luciferase activity was increased, indicating that region 249–437 interacts with the rest of the LBD. Interestingly, this interaction was greatly enhanced when the F domain was present and the lack of the hinge region (construct Gal 383–437) completely abolished the interaction, confirming that this domain is essential in SmFtz-F1 activity. In order to test whether the hinge with helix 1 alone is sufficient to promote the interaction with the remainder of the SmFtz-F1 LBD, we used the Gal 249–409 construct in the same system (Figure 7B). Co-transfected with VP16 alone, this construct displayed transcriptional activity which again corresponds to the fortuitous autonomous transactivation function found in the helix H1. Co-transfection of this construct with VP16 410–632 resulted in a strong increase in luciferase activity compared with Gal 249–409 alone, indicating that the region encompassing the hinge and helix 1 strongly interacts with region 409–628 of the SmFtz-F1 LBD. This effect is specific to helix 1, since we determined that the hinge region alone does not interact with the rest of the LBD (results not shown). Surprisingly, when the F domain was added (construct VP16 410–731), the interaction between the two fusion proteins was weaker. In an identical manner to Gal 383–437, the lack of the hinge domain (construct Gal 383–409) completely prevented interaction. Finally, mutations of Asp-398, Leu-401 and Val-402 (Figure 6A) also prevent interaction between helix 1 and the rest of the LBD (results not shown). Thus the results of the LBD DS assay confirm that the region containing the hinge and helix 1 of the LBD interacts with the remainder of the helical structure of the SmFtz-F1 LBD, and that this interaction implicates residues that are in part involved in the recruitment of fortuitous co-activators when this region is used on its own in transactivation assays (construct Gal-E6; see Figure 4).

The SmFtz-F1 I-box

In addition to observations concerning the N-terminal part of the LBD we were also interested in the C-terminal part of this domain, and more specifically in helices H9 and H10, a region which corresponds to the I-box. The I-box was first identified as a transferable dimerization domain sufficient to control the identity of dimer partners in RXR–RAR and RXR–TR (thyroid hormone receptor) heterodimers [42]. Lee et al. [43] dissected this domain and identified critical residues responsible for heterodimerization in an 11-amino-acid sub-region called heptad 9. More recently, Liu et al. [38] showed that heptad 9 is conserved among Ftz-F1 members, and suggested a role for the I-box in the recruitment of SRC-1 for zebrafish Ftz-F1 α . Our results for the transcriptional activity of the Δ I-box construct (Figure 4) were identical with those obtained with the zebrafish FF1 α mutant [38]. We therefore next asked whether the I-box really influences receptor activity by mediating cofactor recruitment, or whether the lack of transactivation activity of the Δ I-box construct was the result of an overall structural modification or instability of the LBD. For this purpose, we noticed that Gal- Δ E4, which lacks only eight amino acids of which four belong to heptad 9, is inactive (Figures 4 and 5B), suggesting that this stretch of amino acids was vital to the function of heptad 9. We therefore decided to mutate residues in this region and tested mutants for their transcriptional activity and capacity to interact with SRC1. Mutated residues and corresponding constructs are depicted in Figure 8(A), and Figure 8(B) shows the spatial organization of each residue within helix H10. The first two mutants Gal-DKF/AAA and Gal-KLI/AAA

Figure 8 Loss of secondary structure probably provokes the transcriptional inactivity of the I-box mutants

(A) The sequence corresponding to amino acids 579–603 of SmFz-F1 is represented. Residues that are deleted from the Gal- Δ E4 construct are in bold. Residues corresponding to heptad 9 are boxed. Arrowheads show predicted exposed residues of the deleted sequence in the Gal- Δ E4 construct. All the mutations and the corresponding constructs are represented below the sequence. (B) Details of the interactions of helix H10 in structural stabilization. For residues of the DKF motif, contacts are formed between residues of helices H9 and the loop connecting H7 and H8 (L7–8). The residues of the KLI motif stabilize helices H8, H9 and the loop L7–8. (C) Transcriptional activity of the I-box mutants. CV-1 cells were transfected with 200 ng of pFR-Luc reporter vector and 800 ng of Gal4 expression vector. The results are expressed as the percentage of normalized luciferase activity with the activity of the Gal-SmFF1 construct taken as 100%. (D) Interaction between I-box mutants and SRC1RID. CV-1 cells were transfected with 200 ng of pFR-LUC reporter plasmid, 200 ng of Gal4 expression vector and 800 ng of VP16-SRC1RID expression vector. Results are expressed as the fold activity relative to the basal reporter activity obtained by co-transfection with plasmid encoding GAL4 and VP16 alone.

show strongly reduced transactivation activities (Figure 8C), and do not interact with SRC-1 in the two-hybrid assay (Figure 8D). The KLI (Lys-589/Leu-590/Ile-591) motif, in which all the residues are buried, is involved in the maintenance of the LBD architecture, particularly Leu-590 (Figure 8B). Consequently, mutation of this motif most likely perturbs the 3D structure of SmFz-F1, and it is thus not very likely that it is involved in the recruitment of SRC-1. On the other hand, the 3D model reveals that, in the DKF motif (Asp-585/Lys-586/Phe-587), Lys-586 and Phe-587 are mainly involved in the maintenance of the 3D structure, and Asp-585 is exposed (Figure 8B). Thus, in order to test an eventual role of exposed residues in SRC-1 recruitment, we made point mutations of the DKF motif, where each amino acid was replaced by alanine (Figure 8A). As predicted, the K586A mutant shows a strongly reduced luciferase activity (approx. 70% compared with the wild-type; Figure 8C) and an impaired capacity for interaction with SRC-1 (Figure 8D). Interestingly, mutant D585A transactivated significantly (40%) less well than Gal-SmFF1 (Figure 8C), but interacted perfectly with SRC-1 (Figure 8D). This result was unexpected, and clearly indicates that Asp-585 is strongly involved in SmFz-F1 transcriptional activity, perhaps by recruiting alternative cofactors. The F587A mutant exhibited functions similar to those of the wild-type protein.

We also investigated the functional importance of the last two exposed residues of the motif, Asn-588 and Asn-592, by alanine substitution (Figure 8A). Both mutants (N588A and N592A) transactivated the reporter gene and interacted with SRC-1 similarly to Gal-SmFF1 (Figure 8C and results not shown). Thus we found that none of the tested exposed residues are involved in recruitment of SRC-1, but we have identified the exposed Asp-585 as being essential for receptor transcriptional activity, confirming the functional importance of the I-box as an interaction interface. Our results clearly suggest that the lack of activity for the Gal- Δ Ibox and Gal- Δ E4 constructs correlates with a perturbation of the 3D structure, since the mutation of only one buried residue (Lys-586) can induce an identical effect on receptor activity (compare the induced luciferase activity of Gal- Δ E4 in Figure 4 with Gal-K586A in Figure 8C).

LBP: is anybody there?

According to the three models (SmFz-F1rar, SmFz-F1ecr and SmFz-F1lhr), the SmFTZ-F1 LBD exhibits a large hydrophobic cavity (determined to be 689, 690 and 777 Å³ respectively for the three models), except for a few hydrophilic regions formed by polar and charged residues (Arg-448, Asp-485 and Arg-489;

Figure 9 Detailed view of the LBP showing the two arginine residues sealing the pocket

Helices 3 and 5 are indicated, as well as lateral chains of the two arginine residues. 25-Hydroxycholesterol was placed into the pocket to indicate the position and relative dimensions of a putative ligand.

Figure 9). The SmFtz-F1_{rar} and SmFtz-F1_{EcR} models adopt the common agonist fold observed in other nuclear-receptor LBDs [8]. The mouse LRH-1 structure exhibits an unusual active conformation, which seems to be maintained by an extended helix 2 [4]. Moreover, in its large pocket, the conserved Arg-412 in H5, which in human RXR α is involved in the anchoring of the 9-*cis*-retinoic acid carboxylate moiety [44], forms a salt bridge with Asp-408. A similar salt bridge is observed in the SmFtz-F1_{lrh} LBP between Arg-489 and Asp-485, but the poor residue conservation in helix 2 suggests that SmFtz-F1 probably does not adopt an active conformation without a ligand.

Arg-448 is located in helix 3, and might participate in polar interactions with the ligand. Because of these structural features, we decided to test whether SmFtz-F1 might be a ligand-dependent nuclear receptor by using a series of Gal-SmFF1 mutants in co-transfection experiments. Both arginine residues (Arg-448 and Arg-489) were replaced by alanine to abolish the charge. Arg-448 was also replaced by threonine to mimic the SF-1/LRH-1 sequence. We decided to replace Ala-445 located near the ligand-anchoring site as observed in the RAR-9-*cis*-retinoic acid complex with bulkier residues (valine and phenylalanine) in order to interfere with the binding of a putative ligand without perturbing the fold. We also mutated residues Cys-486 and Ala-604 and replaced them with valine. The results in Figure 10(A) show that none of these seven mutants significantly inhibit SmFtz-F1 transcriptional activity. We next decided to use several double mutants, and interestingly the double mutant containing both arginine residues mutated to alanine (Gal-R448A/R489A) shows 50% reduced luciferase activity compared with the wild-type (Figure 10B). We next tested the interaction of this mutant with SRC-1 (Figure 10C), and found that the interaction is reduced by approx. 75%, strongly suggesting that these two buried residues are essential for AF-2 activity. Our results show that two arginine residues, of which the side chains point into the LBP, are essential for receptor AF-2-dependent activity, and we suggest that these residues might be implicated in binding of a putative endogenous ligand. SmFtz-F1 may therefore be a ligand-dependent nuclear receptor.

DISCUSSION

In the present study, we have carried out a systematic analysis of the D, E and F domains of the nuclear receptor SmFtz-F1. Our first

Figure 10 LBP mutants suggest that SmFtz-F1 might be a ligand-dependent nuclear receptor

(A) Transcriptional activity of single amino acid LBP mutants. CV-1 cells were transfected with 200 ng of pFR-Luc reporter vector and 800 ng of Gal4 expression vector. Results are expressed as the percentage of normalized luciferase activity with the activity of the Gal-SmFF1 construct taken as 100%. (B) The double arginine LBP mutant shows impaired transcriptional activity compared with the Gal-SmFF1 construct. CV-1 cells were transfected with 200 ng of pFR-Luc reporter vector and 800 ng of Gal4 expression vector. Results are expressed as described for relevant previous Figures. (C) The R448A/R489A mutant strongly inhibits interaction of SmFtz-F1 with SRC1_{IRID}. CV-1 cells were transfected with 200 ng of pFR-Luc reporter vector and 800 ng of Gal4 expression vector. Results are expressed as the fold activity relative to the basal reporter activity obtained by co-transfection with plasmid encoding GAL4 and VP16 alone.

paper concerning this receptor demonstrated that, despite a relatively weak sequence conservation, SmFtz-F1 conserves the basic functional characteristics of the subfamily [23]. In the present paper, our results show that the transcriptional activity of SmFtz-F1 is dependent on the D, E and F domains and that, like other nuclear receptors, the recruitment of co-activators is

mediated through the AF-2-AD domain. Despite the conservation of the α -helical sandwich fold, our results highlight the functional importance of two poorly characterized domains of nuclear receptors, the hinge region and the F domain, and mutational analysis of the predicted LBP seems to suggest that SmFtz-F1 might be a ligand-dependent nuclear receptor.

Unlike other members of the Ftz-F1 subfamily, SmFtz-F1 possesses a long F domain of unknown function. Our transfection experiments revealed that the F domain is essential for the full activity of the receptor, but the localization of a particular functional motif proved elusive. In particular, the transactivation levels and capacity for interaction with SRC1 of deleted proteins did not allow us to delineate a precise portion of the peptide crucial in the role of the F domain, but the results suggest rather that the entire domain is involved in maintaining the functional integrity of SmFtz-F1. As we previously noted, not all nuclear receptors contain an F domain. In the case of steroid receptors, it most often corresponds to a short C-terminal extension that has been shown to influence ligand binding and transactivational properties of different receptors [45,46]. The ER (oestrogen receptor) F domain is longer, and Nichols et al. [47] clearly demonstrated the variable position of this domain, along with helix 12, according to the presence of an agonist or an antagonist, therefore probably influencing interaction with cofactors. Thus, in this regard, it is not surprising that our two-hybrid assay clearly showed that the F domain strongly contributes to the AF-2-dependent activity of the receptor (Figure 3B). It seems evident that the position of at least part of the F domain is dictated by the placement of helix 12, but apparently none of the AF-2 residues are hidden by this large C-terminal region, as is probably the case for hepatocyte nuclear factor 4 [48]. One possibility is that it plays the same role as a ligand by maintaining helix 12 in an active conformation, tightly folded against the LBD. It could also simply constitute an additional interaction surface either required for the AF-2 activity of the receptor or allowing interaction with specific schistosome cofactors implicated in a more subtle transcriptional regulation. Further investigations to determine whether potential contact between the LBD and the F domain exists, and the nature of the residues implicated in this interaction, are clearly required. An X-ray crystal structure of the LBD with its F domain would also be very helpful to understand this probably novel mechanism of transactivational regulation.

Little is known about the function of the hinge region. The variable length and poor conservation of the D domain in different receptors might reflect a lack of functional conservation. For example, mutagenesis of the hinge reduces the transcriptional activity of the glucocorticoid receptor, without affecting its ligand-binding capacity [49], but has no effect on the function of the ER [50]. Recent reports have underlined the importance of this region, notably in maintaining the functional conformation of nuclear-receptor LBDs [38,39]. Our results concerning the D domain of SmFtz-F1 seem also to argue for an important conformational role of this domain. First, the LBD dynamic functional assay required the presence of the D domain of SmFtz-F1. Moreover, the ablation of the SmFtz-F1 hinge region completely abolished the transcriptional activity, and particularly the interaction with SRC-1, confirming an essential role in promoting an active conformation of the LBD. Adding the 90 C-terminal amino acids of the hinge restored a normal interaction with SRC-1 (construct Gal- Δ D3; see Figure 5A), but was insufficient to retrieve full transcriptional activity (construct Gal- Δ D3; see Figure 4). Thus this result also suggests that a positive regulatory sequence could be present between amino acids 249–295 in the hinge region. Indeed, several authors have reported that the hinge region can integrate regulation signals, and that post-translational modifica-

tions can regulate receptor activity. For example, the acetylation of the ER hinge region by p300 suppressed ligand sensitivity [51]. In the case of SF-1, phosphorylation by a MAPK of a single serine residue of the D domain appears to promote an active conformation and increase the transcriptional activity of the receptor [22]. Phosphorylation of a serine residue in the hinge region of mouse ER [52] and human androgen receptor [53] have also been described. Even though, unlike SF-1, no MAPK kinase consensus sequence is present in the hinge region of SmFtz-F1, several serine residues are predicted to be potential kinase targets, notably for casein kinase II (not shown), indicating that phosphorylation might regulate SmFtz-F1 transcriptional activity. Another possibility is that region 249–295 constitutes a binding surface for co-activators. For example, Onate et al. [54] have demonstrated that SRC-1 harbours two receptor-binding regions and interacts with both the N-terminal A/B domain and the C-terminal D/E domains of PR. It is possible therefore that the full activity of SmFtz-F1 also requires two interacting surfaces, one represented by the well-characterized AF-2-AD and the other being present in the hinge, between residues 249–295. This hypothesis will be tested in the mammalian two-hybrid system using constructs that are further deleted or mutated in this region. Interestingly, we have also noticed that the hinge alone significantly transactivates the reporter gene (construct Gal- Δ E5 in Figure 4). This result is in favour of the presence of a regulatory motif in the hinge region of SmFtz-F1.

It clearly appears that, for many receptors, the flanking domains of the LBD, the D and F domains, are essential for nuclear receptor function. Nevertheless, structural data are only available at present for isolated nuclear receptor DBDs and LBDs. These crystal structures have allowed us to understand essential conserved mechanisms such as DNA binding, interaction with cofactors and ligand-dependent transactivation, but to completely understand how nuclear receptors work we need to determine the structural relationship between the different receptor domains as reported recently for human RXR α [55].

Among the different families of nuclear receptors, the sequence identity of the E domain is quite low. Despite this divergence, the 3D structures of the LBDs are very similar [8], allowing the development of computer-assisted modelling based on previously established structures. We used this strategy to construct a model of the SmFtz-F1 LBD that permitted rational analysis of this domain. Notably, we have based our mutations and deletions on the 12-helix structure, with the aim of avoiding disruptions in this structure which might lead to artefactual changes in transcriptional activity and/or cofactor binding. Our mutational analysis has strongly suggested that the activation function found in helix 1 is fortuitous, because this region is, as in other nuclear receptors, implicated in the architecture of the LBD. Indeed, recent reports have highlighted the role of helix 1 in the structural stabilization of various nuclear receptors. Pissios et al. [24] first reported a ligand-dependent interaction between helix 1 and the remainder of the TR LBD. They proposed that this interaction stabilizes the overall structure of the LBD. This was then confirmed when helix 1 of the TR LBD was found to be essential for co-activation by PGC-1 (peroxisome-proliferator-activated receptor- γ co-activator 1) and GRIP-1 (glucocorticoid receptor-interacting protein 1), indicating that this region maintains LBD structural integrity [56]. Holt et al. [41] have extended this mechanism to an orphan receptor [NGF1-B (nerve growth factor-induced-B)] and identified critical residues implicated in intramolecular interactions between helices 1 and 8, suggesting that it is likely that the stabilization function of helix 1 is an evolutionarily conserved feature of nuclear receptors. We used a similar approach and showed that helix 1 interacts with the rest of

the LBD of SmFtz-F1, and that this interaction implicates residues that are in part involved in the recruitment of a fortuitous co-activator in the case of the Gal- Δ E6 construct, for example. The strongest interaction was obtained when the region connecting H1 and H3 (amino acids 409–437) is linked to helix H3 in the VP16 410–632 construct (Figure 7B) rather than when attached to helix H1 in the Gal 249–437 construct. This result highlights the role of this region that probably stabilizes the helical structure of the LBD, allowing a stronger interaction with helix H1. Recent results have already shown that this same region influences the overall stability of the thyroid receptor LBD [57]. Moreover, in the case of SmFtz-F1, the F domain possesses the particularity of modulating the interaction between helix H1 and the remainder of the LBD, strengthening the hypothesis that this domain is involved in LBD functions. However, the F domain displayed opposite effects on the interaction in the two different types of construct. This may reflect a differential interaction of the F domain with the remaining structure when the H2 region is present on one or the other interacting constructs.

We finally attempted to determine whether SmFtz-F1 could be a ligand-dependent nuclear receptor. SmFtz-F1 belongs to the SF-1/LRH-1 subfamily that only contains orphan receptors with no identified ligands. The crystal structure of the LRH-1 LBD recently provided evidence for ligand-independent transcription activity of members of this subfamily of nuclear receptors [4]. The authors have proposed that the active conformation of the receptor was due to a subfamily-specific helix 2, tightly packed against helix H3, that stabilizes the LBD in the absence of a ligand [4]. The sequence identity between the E domain of SmFtz-F1 and human SF-1 is only 20%, and the region corresponding to helix H2 is poorly conserved and will probably present a different 3D conformation. This raises the question of the regulation of SmFtz-F1 activity and the existence of an agonist ligand for this nuclear receptor. Our results clearly demonstrate the functional importance of two arginine residues, the side chains of which are directed into the LBP. The capacity of the double arginine mutant to interact with SRC-1 is strongly diminished, suggesting that the folding of AF-2 is not optimal in this mutant. This might be due to a disruption in the structure of the LBD, but we consider this unlikely, since the side chains of the two arginine residues are free within the LBP. We therefore favour the hypothesis of the presence of a fortuitous ligand in the active SmFtz-F1 LBD expressed in CV-1 cells. We are currently pursuing these studies with the construction of further multiple LBP mutants. Moreover, we envisage X-ray crystallographic analysis of the SmFtz-F1 LBD in order to obtain a definitive answer to this important question.

Overall, the structural and functional properties of SmFtz-F1 show some very marked differences compared with its mammalian counterparts, and illustrate the way in which the evolution of nuclear receptor functions may contribute to, or at least follow, species diversification. The latest expressed sequence tag data for *S. mansoni* [58] suggest that schistosomes have a relatively limited complement of nuclear receptors, in marked contrast with *C. elegans*. It will be interesting to determine whether this low number of nuclear receptors is compensated by a greater flexibility in their functional capacity.

We thank Dr B. W. O'Malley for providing human SRC-1 expression vector, Dr J. L. Jameson for VP16 expression vectors and Dr Philippe Lefebvre (INSERM U459, Faculté de Médecine Henri Warembourg, Lille, France) for the pVP16 vector. This work was supported by the Institut National de la Santé et de la Recherche Médicale (U547), the Institut Pasteur de Lille, the Centre National de la Recherche Scientifique and the Microbiology program of the Ministère de l'Éducation Nationale, de la Recherche et de la Technologie (MENRT). B.B. is supported by the MENRT. F.O. is supported by INSERM-Region Nord Pas de Calais.

REFERENCES

- Gronemeyer, H. and Laudet, V. (1995) Transcription factors 3: nuclear receptors. *Protein Profile* **2**, 1173–1308.
- Giguere, V. (1999) Orphan nuclear receptors: from gene to function. *Endocr. Rev.* **20**, 689–725.
- Kallen, J. A., Schlaeppli, J. M., Bitsch, F., Geisse, S., Geiser, M., Delhon, I. and Fournier, B. (2002) X-ray structure of the hROR α LBD at 1.63 Å: structural and functional data that cholesterol or a cholesterol derivative is the natural ligand of ROR α . *Structure (Cambridge)* **10**, 1697–1707.
- Sablín, E. P., Krylova, I. N., Fletterick, R. J. and Ingraham, H. A. (2003) Structural basis for ligand-independent activation of the orphan nuclear receptor LRH-1. *Mol. Cell* **11**, 1575–1585.
- Escriva, H., Safi, R., Hanni, C., Langlois, M. C., Saumitou-Laprade, P., Stehelin, D., Capron, A., Pierce, R. and Laudet, V. (1997) Ligand binding was acquired during evolution of nuclear receptors. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 6803–6808.
- Renaud, J. P., Rochel, N., Ruff, M., Vivat, V., Chambon, P., Gronemeyer, H. and Moras, D. (1995) Crystal structure of the RAR- γ ligand-binding domain bound to all-trans retinoic acid. *Nature (London)* **378**, 681–689.
- Nolte, R. T., Wisely, G. B., Westin, S., Cobb, J. E., Lambert, M. H., Kurokawa, R., Rosenfeld, M. G., Willson, T. M., Glass, C. K. and Milburn, M. V. (1998) Ligand binding and co-activator assembly of the peroxisome proliferator-activated receptor- γ . *Nature (London)* **395**, 137–143.
- Wurtz, J. M., Bourguet, W., Renaud, J. P., Vivat, V., Chambon, P., Moras, D. and Gronemeyer, H. (1996) A canonical structure for the ligand-binding domain of nuclear receptors. *Nat. Struct. Biol.* **3**, 206.
- Horlein, A. J., Naar, A. M., Heinzel, T., Torchia, J., Gloss, B., Kurokawa, R., Ryan, A., Kamei, Y., Soderstrom, M., Glass, C. K. et al. (1995) Ligand-independent repression by the thyroid hormone receptor mediated by a nuclear receptor co-repressor. *Nature (London)* **377**, 397–404.
- Laudet, V., Auwerx, J., Gustafsson, J. A. and Wahli, W. (1999) A unified nomenclature system for the nuclear receptor superfamily. *Cell* **97**, 161–163.
- Ueda, H., Sonoda, S., Brown, J. L., Scott, M. P. and Wu, C. (1990) A sequence-specific DNA-binding protein that activates fushi tarazu segmentation gene expression. *Genes Dev.* **4**, 624–635.
- Yu, Y., Li, W., Su, K., Yussa, M., Han, W., Perrimon, N. and Pick, L. (1997) The nuclear hormone receptor Ftz-F1 is a cofactor for the *Drosophila* homeodomain protein Ftz. *Nature (London)* **385**, 552–555.
- Asahina, M., Ishihara, T., Jindra, M., Kohara, Y., Katsura, I. and Hirose, S. (2000) The conserved nuclear receptor Ftz-F1 is required for embryogenesis, moulting and reproduction in *Caenorhabditis elegans*. *Genes Cells* **5**, 711–723.
- Honda, S., Morohashi, K., Nomura, M., Takeya, H., Kitajima, M. and Omura, T. (1993) Ad4BP regulating steroidogenic P-450 gene is a member of steroid hormone receptor superfamily. *J. Biol. Chem.* **268**, 7494–7502.
- Parker, K. L. and Schimmer, B. P. (1997) Steroidogenic factor 1: a key determinant of endocrine development and function. *Endocr. Rev.* **18**, 361–377.
- Tugwood, J. D., Isserman, I. and Green, S. (1991) Mouse liver receptor homologous protein (LRH-1) mRNA. GenBank® Nucleotide Sequence Database, entry M81385.
- Galarneau, L., Pare, J. F., Allard, D., Hamel, D., Levesque, L., Tugwood, J. D., Green, S. and Belanger, L. (1996) The α 1-fetoprotein locus is activated by a nuclear receptor of the *Drosophila* FTZ-F1 family. *Mol. Cell. Biol.* **16**, 3853–3865.
- Becker-Andre, M., Andre, E. and DeLamarier, J. F. (1993) Identification of nuclear receptor mRNAs by RT-PCR amplification of conserved zinc-finger motif sequences. *Biochem. Biophys. Res. Commun.* **194**, 1371–1379.
- Lin, W., Wang, H. W., Sum, C., Liu, D., Hew, C. L. and Chung, B. (2000) Zebrafish ftz-f1 gene has two promoters, is alternatively spliced, and is expressed in digestive organs. *Biochem. J.* **348**, 439–446.
- Lu, T. T., Makishima, M., Repa, J. J., Schoonjans, K., Kerr, T. A., Auwerx, J. and Mangelsdorf, D. J. (2000) Molecular basis for feedback regulation of bile acid synthesis by nuclear receptors. *Mol. Cell* **6**, 507–515.
- Mellon, S. H. and Bair, S. R. (1998) 25-Hydroxycholesterol is not a ligand for the orphan nuclear receptor steroidogenic factor-1 (SF-1). *Endocrinology* **139**, 3026–3029.
- Desclozeaux, M., Krylova, I. N., Horn, F., Fletterick, R. J. and Ingraham, H. A. (2002) Phosphorylation and intramolecular stabilization of the ligand binding domain in the nuclear receptor steroidogenic factor 1. *Mol. Cell. Biol.* **22**, 7193–7203.
- De Mendonca, R. L., Bouton, D., Bertin, B., Escriva, H., Noel, C., Vanacker, J. M., Cornette, J., Laudet, V. and Pierce, R. J. (2002) A functionally conserved member of the FTZ-F1 nuclear receptor family from *Schistosoma mansoni*. *Eur. J. Biochem.* **269**, 5700–5711.
- Pissios, P., Tzameli, I., Kushner, P. and Moore, D. D. (2000) Dynamic stabilization of nuclear receptor ligand binding domains by hormone or corepressor binding. *Mol. Cell* **6**, 245–253.

- 25 Carey, M., Kakidani, H., Leatherwood, J., Mostashari, F. and Ptashne, M. (1989) An amino-terminal fragment of GAL4 binds DNA as a dimer. *J. Mol. Biol.* **209**, 423–432
- 26 Yao, T. P., Ku, G., Zhou, N., Scully, R. and Livingston, D. M. (1996) The nuclear hormone receptor coactivator SRC-1 is a specific target of p300. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 10626–10631
- 27 Sambrook, J. and Russel, D. W. (2001) *Molecular Cloning: A Laboratory Manual*, 3rd edn, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- 28 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 (London)* **426**, 91–96
- 29 Sali, A. and Blundell, T. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815
- 30 Laskowski, R. A., MacArthur, M. W., Moss, D. S. and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structure. *J. Appl. Cryst.* **26**, 283–291
- 31 Hendlich, M., Lackner, P., Weitckus, S., Floeckner, H., Froschauer, R., Gottsbacher, K., Casari, G. and Sippl, M. (1990) Identification of native protein folds amongst a large number of incorrect models. The calculation of low energy conformations from potentials of mean force. *J. Mol. Biol.* **216**, 167–180
- 32 Leo, C. and Chen, J. D. (2000) The SRC family of nuclear receptor coactivators. *Gene* **245**, 1–11
- 33 Ito, M., Yu, R. N. and Jameson, J. L. (1998) Steroidogenic factor-1 contains a carboxy-terminal transcriptional activation domain that interacts with steroid receptor coactivator-1. *Mol. Endocrinol.* **12**, 290–301
- 34 Durand, B., Saunders, M., Gaudon, C., Roy, B., Losson, R. and Chambon, P. (1994) Activation function 2 (AF-2) of retinoic acid receptor and 9-cis retinoic acid receptor: presence of a conserved autonomous constitutive activating domain and influence of the nature of the response element on AF-2 activity. *EMBO J.* **13**, 5370–5382
- 35 Crawford, P. A., Polish, J. A., Ganpule, G. and Sadovsky, Y. (1997) The activation function-2 hexamer of steroidogenic factor-1 is required, but not sufficient for potentiation by SRC-1. *Mol. Endocrinol.* **11**, 1626–1635
- 36 He, B., Kempainen, J. A. and Wilson, E. M. (2000) FXFLF and WXXLF sequences mediate the NH₂-terminal interaction with the ligand binding domain of the androgen receptor. *J. Biol. Chem.* **275**, 22986–22994
- 37 Heery, D. M., Kalkhoven, E., Hoare, S. and Parker, M. G. (1997) A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature (London)* **387**, 733–736
- 38 Liu, D., Chandy, M., Lee, S. K., Le Dren, Y., Ando, H., Xiong, F., Woon Lee, J. and Hew, C. L. (2000) A zebrafish ftz-F1 (Fushi tarazu factor 1) homologue requires multiple subdomains in the D and E regions for its transcriptional activity. *J. Biol. Chem.* **275**, 16758–16766
- 39 Miyamoto, T., Kakizawa, T., Ichikawa, K., Nishio, S., Takeda, T., Suzuki, S., Kaneko, A., Kumagai, M., Mori, J., Yamashita, K. et al. (2001) The role of hinge domain in heterodimerization and specific DNA recognition by nuclear receptors. *Mol. Cell. Endocrinol.* **181**, 229–238
- 40 Rost, B., Sander, C. and Schneider, R. (1994) PHD – an automatic mail server for protein secondary structure prediction. *Comput. Appl. Biosci.* **10**, 53–60
- 41 Holt, J. A., Consler, T. G., Williams, S. P., Ayscue, A. H., Leesnitzer, L. M., Wisely, G. B. and Billin, A. N. (2003) Helix 1/8 interactions influence the activity of nuclear receptor ligand-binding domains. *Mol. Endocrinol.* **17**, 1704–1714
- 42 Perlmann, T., Umesono, K., Rangarajan, P. N., Forman, B. M. and Evans, R. M. (1996) Two distinct dimerization interfaces differentially modulate target gene specificity of nuclear hormone receptors. *Mol. Endocrinol.* **10**, 958–966
- 43 Lee, S. K., Na, S. Y., Kim, H. J., Soh, J., Choi, H. S. and Lee, J. W. (1998) Identification of critical residues for heterodimerization within the ligand-binding domain of retinoid X receptor. *Mol. Endocrinol.* **12**, 325–332
- 44 Egea, P. F., Mitschler, A., Rochel, N., Ruff, M., Chambon, P. and Moras, D. (2000) Crystal structure of the human RXR α ligand-binding domain bound to its natural ligand: 9-cis retinoic acid. *EMBO J.* **19**, 2592–2601
- 45 Xu, J., Nawaz, Z., Tsai, S. Y., Tsai, M. J. and O'Malley, B. W. (1996) The extreme C terminus of progesterone receptor contains a transcriptional repressor domain that functions through a putative corepressor. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 12195–12199
- 46 Zhang, S., Liang, X. and Danielsen, M. (1996) Role of the C terminus of the glucocorticoid receptor in hormone binding and agonist/antagonist discrimination. *Mol. Endocrinol.* **10**, 24–34
- 47 Nichols, M., Rientjes, J. M. and Stewart, A. F. (1998) Different positioning of the ligand-binding domain helix 12 and the F domain of the estrogen receptor accounts for functional differences between agonists and antagonists. *EMBO J.* **17**, 765–773
- 48 Sladek, F. M., Ruse, Jr, M. D., Nepomuceno, L., Huang, S. M. and Stallcup, M. R. (1999) Modulation of transcriptional activation and coactivator interaction by a splicing variation in the F domain of nuclear receptor hepatocyte nuclear factor 4 α 1. *Mol. Cell. Biol.* **19**, 6509–6522
- 49 Giguere, V., Hollenberg, S., Rosenfeld, M. and Evans, R. (1986) Functional domains of the human glucocorticoid receptor. *Cell* **46**, 645–652
- 50 Kumar, V., Green, S., Stack, G., Berry, M., Jin, J. and Chambon, P. (1987) Functional domains of the human estrogen receptor. *Cell* **51**, 941–951
- 51 Wang, C., Fu, M., Angeletti, R. H., Siconolfi-Baez, L., Reutens, A. T., Albanese, C., Lisanti, M. P., Katzenellenbogen, B. S., Kato, S., Hopp, T. et al. (2001) Direct acetylation of the estrogen receptor α hinge region by p300 regulates transactivation and hormone sensitivity. *J. Biol. Chem.* **276**, 18375–18383
- 52 Lahoti, H., White, R., Hoare, S. A., Rahman, D., Pappin, D. J. and Parker, M. G. (1995) Identification of phosphorylation sites in the mouse oestrogen receptor. *J. Steroid. Biochem. Mol. Biol.* **55**, 305–313
- 53 Zhou, Z. X., Kempainen, J. A. and Wilson, E. M. (1995) Identification of three proline-directed phosphorylation sites in the human androgen receptor. *Mol. Endocrinol.* **9**, 605–615
- 54 Onate, S. A., Boonyaratankornkit, V., Spencer, T. E., Tsai, S. Y., Tsai, M. J., Edwards, D. P. and O'Malley, B. W. (1998) The steroid receptor coactivator-1 contains multiple receptor interacting and activation domains that cooperatively enhance the activation function 1 (AF1) and AF-2 domains of steroid receptors. *J. Biol. Chem.* **273**, 12101–12108
- 55 Fischer, H., Dias, S. M., Santos, M. A., Alves, A. C., Zanchin, N., Craievich, A. F., Apriletti, J. W., Baxter, J. D., Webb, P., Neves, F. A. et al. (2003) Low resolution structures of the retinoid X receptor DNA-binding and ligand-binding domains revealed by synchrotron X-ray solution scattering. *J. Biol. Chem.* **278**, 16030–16038
- 56 Wu, Y., Delerive, P., Chin, W. W. and Burris, T. P. (2002) Requirement of helix 1 and the AF-2 domain of the thyroid hormone receptor for coactivation by PGC-1. *J. Biol. Chem.* **277**, 8898–8905
- 57 Huber, B., Desclozeaux, M., West, B., Cunha-Lima, S., Nguyen, H., Baxter, J., Ingraham, H. and Fletterick, R. (2003) Thyroid hormone receptor- β mutations conferring hormone resistance and reduced corepressor release exhibit decreased stability in the N-terminal ligand-binding domain. *Mol. Endocrinol.* **17**, 107–116
- 58 Verjovski-Almeida, S., DeMarco, R., Martins, E., Guimaraes, P., Ojopi, E., Paquola, A., Piazza, J., Nishiyama, M., Kitajima, J., Adamson, R. et al. (2003) Transcriptome analysis of the acelomate human parasite *Schistosoma mansoni*. *Nat. Genet.* **35**, 148–157

Received 25 March 2004/16 April 2004; accepted 23 April 2004

Published as BJ Immediate Publication 23 April 2004, DOI 10.1042/BJ20040489