



Structure-based analysis of the ultraspiracle protein and docking studies of putative ligands

Souphatta Sasorith, Isabelle M.L. Billas, Thomas Iwema, Dino Moras and Jean-Marie Wurtz

Département de Génomique et de Biologie Structurales, Institut de Génétique et de Biologie Moléculaire et Cellulaire, 1, rue Laurent Fries, 67404 Illkirch, France.
billas@titus.u-strasbg.fr

Received 14 August 2002, Accepted 21 October 2002, Published 3 December 2002

This article was featured in the Ecdysone Workshop

Abstract

The ultraspiracle protein (USP) is the insect ortholog of the mammalian retinoid X receptor (RXR). Fundamental questions concern the functional role of USP as the heterodimerization partner of insect nuclear receptors such as the ecdysone receptor. The crystallographic structures of the ligand binding domain of USPs of *Heliothis virescens* and *Drosophila melanogaster* solved recently show that helix 12 is locked in an antagonist conformation raising the question whether USPs could adopt an agonist conformation as observed in RXR α . In order to investigate this hypothesis, a homology model for USP is proposed that allows a structural analysis of the agonist conformation of helix 12 based on the sequence comparison with RXR. For USP, one of the main issues concerns its function and in particular whether its activity is ligand independent or not. The x-ray structures strongly suggest that USP can bind ligands. Putative ligands have therefore been docked in the USP homology model. Juvenile hormones and juvenile hormone analogs were chosen as target ligands for the docking study. The interaction between the ligand and the receptor are examined in terms of the pocket shape as well as in terms of the chemical nature of the residues lining the ligand binding cavity.

Keywords: nuclear receptor, ligand-dependent regulation of transcription, ligand binding domain, ultraspiracle protein, homology modeling, agonist, antagonist, juvenile hormone

Abbreviation:

USP	ultraspiracle protein
RXR	retinoid X receptor
LBD	ligand binding domain
LBP	ligand binding pocket
EcR	ecdysone receptor
9-cis RA	9-cis retinoic acid
JH	juvenile hormone

Introduction

The ultraspiracle protein (USP) belongs to the superfamily of nuclear receptors, which are ligand-inducible transcription regulators. USP is the insect ortholog of the vertebrate retinoid X receptor (RXR) (Oro et al., 1990). Like RXR, USP heterodimerizes with nuclear receptors to form active receptor complexes. In particular, USP heterodimerizes with the ecdysone receptor (EcR), the receptor of ecdysteroids which are insect steroid hormones that control insect development, reproduction, molting and metamorphosis (for a review, see Riddiford et al., 2001)). The heterodimerization of EcR with USP is necessary for high affinity binding of ecdysteroids to EcR and transcriptional activity (Yao et al., 1993). In addition to EcR, USP dimerizes with HR38 (Sutherland

et al., 1995) and with HR78 (Hirai et al., 2002), two other insect nuclear receptors. A fundamental question concerns whether the activity of USP is mediated through ligand binding or not, as USP is an orphan receptor for which no endogenous ligand has been established unambiguously. For RXR, agonistic ligands like the 9-cis retinoic acid (9-cis RA), the dodecosahexaenoic acid or the phytanic acid (McCarty, 2001) are known to bind to RXR and modulate the activity of the RXR dimeric partner. For USP, juvenile hormones have been proposed as endogenous ligands that directly modulate the activity of the EcR/USP complex (Jones & Sharp, 1997; Jones et al., 2001). In fact, juvenile hormones, which belong to the family of terpenes as do retinoids, have been shown to prevent metamorphosis by modulating the ecdysteroid action at the outset of the ecdysteroid rise for the molt (Truman & Riddiford, 2002).

Recently, the crystallographic structures of the ligand binding domain (LBD) of USPs from the lepidopteran *Heliothis virescens* (hvUSP) and the dipteran *Drosophila melanogaster* (dmUSP) have been solved (Billas et al., 2001; Clayton et al., 2001). The structures show the canonical fold of nuclear receptor ligand binding domains. The helix H12 at the carboxy-terminal end of the domain, which bears the ligand-dependent AF-2 activation function, adopts a position similar to that observed for other nuclear receptors complexed with antagonistic ligands. However, in the case of USP, it is not a ligand that is responsible for the antagonist conformation of H12, but rather a secondary structural element of the protein itself in which the loop connecting helices 1 and 3 precludes the canonical agonist conformation, and stabilizes the antagonist position of H12. The high sequence conservation of this stretch suggests its functional relevance and its specificity gained during the evolution of highly derived holometabolous insects. The ligand binding pocket (LBP) is large and mostly hydrophobic. It contains a copurified lipid, which originates from *Escherichia coli* used as the protein expression host. Further experimental evidence would be necessary in order to draw conclusions about the physiological significance of this phospholipid as a USP ligand. On the other hand, the high sequence conservation of the residues of USP and RXR, and particularly those lining the ligand binding pocket, tends to suggest that ligands of similar nature and size as those observed for RXR might also exist for USP.

We present here a structure-based analysis of the ultraspiracle protein. The purpose of this study is two-fold. First, we propose a homology model for hvUSP, which includes features of the crystal structures of hvUSP and hsRXR α . It is characterized by the absence of the lipid from the LBP and the canonical agonist conformation of helix H12. These peculiar aspects of the model, and in particular the agonist conformation of H12 are examined structurally on the basis of the sequence homology with RXRs. Secondly, this model was used as a template for the docking of juvenile hormones and their analogs using state-of-the-art biomolecular calculations. The ligand conformations and configurations in the LBP are discussed in terms of energy and configuration. Furthermore, from the shape of the cavity and the chemical nature of the residues lining the LBP, interactions between the ligand functional groups and these residues are examined.

Materials and Methods

Generation of protein models

A homology model of hvUSP was generated according to the crystal structure of hsRXR α as observed in the hsRXR α in complex with its ligand 9-cis RA using the software Modeller (Sali & Blundell, 1993) with standard parameters. Before generating this homology model, the loop connecting helices H1 and H3 (loop 1-3) in the hsRXR α was constructed using the two monomers of apo RXR α found in the asymmetric unit which both present a different and incomplete loop 1-3. The combination of these two structures suggests a complete trace for the loop 1-3 that was further used in order to generate a full USP model with H12 in an agonist position. The second model is based on the hvUSP crystal structure. The third homology model is a chimeric structure, which combines

features observed in the crystal structures of both hsRXR α and hvUSP.

In order to generate the hybrid molecule, the hvUSP crystal structure and the hvUSP RXR-like model were first superimposed using Lsq-man of the O package (version 6) (Jones et al., 1991). The C-terminal part of hvUSP was replaced by that of the RXR template, resulting in an hvUSP model with the loop 11-12 and H12 in the canonical agonist conformation. Additional adjustments had to be done after removal of the phospholipid. In fact, the helix H3 had to be readjusted until it formed van der Waals contacts with H11. The van der Waals interaction is essentially an attractive force that is spontaneously generated between two atoms in their ground state due to the instantaneous polarization of their electronic cloud. The readjustment of H3 was achieved by bending the N-terminal region (Phe 242 to Gln 256) of H3 towards the core of the protein and leaving the C-terminal region unchanged. This hybrid molecule was used as a template for generating the LBD of USPs from different insect types.

The quality of the models was evaluated by both Procheck (Laskowski et al., 1993) and ProsaII (Hendlich et al., 1990). Procheck indicates the percentage of residues located in the favored regions of the Ramachandran plot. This plot gives the main chain conformation as pairs of ϕ and ψ dihedral angles for each residue in the protein. The stereochemical parameters show that more than 97% of the residues of all models have the (ϕ, ψ) dihedral angles in the most favored and the allowed regions of the Ramachandran plot as expected for a good model. The second quality check is based on the Z-score. This index indicates by how much the generated structure differs from the noise. For a given three-dimensional protein structure, the residues are approximated as spheres and randomly distributed along the polypeptide chain. For each distribution of residues, a pairwise energy is calculated. A mean energy and a standard deviation are then evaluated for these distributions. The Z-score of a given structure gives the distance from the mean energy in standard deviation units. The score of native protein folds are usually negative and in characteristic range (-7 to -10 for experimental nuclear receptor refined crystal structures). The Z-score of the hybrid model indicates a good model quality.

Volume calculations

The volume of the pockets and ligands were calculated using GRASP (Nicholls et al., 1991).

Docking experiments and analysis of protein-ligand interactions

The ligands including all hydrogen atoms were built and minimized with the Quanta / Charmm package (Accelrys) (Brooks et al., 1983). The residues inside the ligand binding cavity were identified by the software Voidoo (Jones et al., 1991) and their side chains were positioned using the O rotamer library with an adapted rotamer library from Dunbrack (Dunbrack & Karplus, 1993). Energy minimization of the hybrid model with the all atom force field was performed using the Powell algorithm for 1000 steps. The parameters for this minimization were the following: a dielectric constant of 4, a minimization step of 0.02 and the non-bonded interaction list was set to 15 Å.

The docking procedure proceeds in several steps. First, the protein-ligand complex is generated using the Gold package (Jones

et al., 1997) without constraints between the ligand and specific amino acids of the pocket. The algorithm exhaustively searches the entire rotational and translational space of the ligand with respect to the receptor. The flexibility of the ligand is given by dihedral angle variations. The various solutions are evaluated by a score, which is equivalent to the absolute value of the total energy of the ligand in the protein environment. Because Gold considers the side chains of the amino acids in the pocket as rigid, we used Charmm in an all atom force field to relax their conformation. The energy of the complex is minimized in three successive steps using the Powell algorithm. First, the hydrogen atoms of the complex are considered and their positions are allowed to relax. Then, for a fixed protein conformation, the ligand is considered and its position and conformation are refined by energy minimization. Finally, the ligand in the conformation found by energy minimization together with the side chains of residues located at a radius of less than 10 Å from the ligand are considered and their positions are allowed to relax. In order to score the different solutions of ligand-protein complex, the total interaction energy comprising the van der Waals and the electrostatic contributions are taken into account. The deformation energy term as well as the solvation term is neglected in a first approximation since they contribute rather equally for the very similar ligands considered in this study.

Results

Structure-based sequence alignment

The alignment of the sequences of the USP and RXR LBDs is presented in Figure 1. The USP LBD sequences are subdivided into two groups corresponding to the insect orders Lepidoptera and Diptera. The RXR LBD sequences comprise the RXR sequences of arthropods other than the Lepidoptera and Diptera (*Apis mellifera*, *Locusta migratoria*, *Tenebrio molitor*, *Uca pugillator*, *Amblyomma americanum*) and of vertebrate members of the three isotypes α , β and γ .

The RXR LBD sequences of vertebrates and arthropods are highly conserved having more than 60 % identity. Similarly, a high sequence conservation is seen within the lepidopteran family, with above 80 % identity, while the dipteran USP LBD sequences are less well conserved inside the order, having between 47 and 74 % identity. Altogether, the USP LBD sequences are rather well conserved with respect to those of RXR LBDs, with between 42 and 51 % identity. The crystal structures of hsRXR α and hvUSP, 11 helices and a β -sheet, confirm the secondary structure prediction using the canonical structure of nuclear receptor LBDs (Wurtz et al., 1996). The alignment also highlights an insertion between helices H6 and H7, which seems to be conserved for RXR sequences of various types of fishes. Structurally, the major difference between the RXR and USP sequences is a long insertion between helix H5 and the β -sheet observed for all USPs, but absent in all vertebrate and arthropod RXRs. Most of the conserved residues between USP and RXR sequences are located in the helices, in particular those forming the core of the LBD as well as the signature region which encompasses the C-terminal region of H3, the loop 3-4 and most of helix H4 (Wurtz et al., 1996). In particular, in H4, the aspartate residue (Asp278 for hvUSP) found in the motif DQVI of RXRs, is strictly conserved among all RXRs and USPs. This residue plays a

structural role in the stabilization of the loop between helices H8 and H9.

If we consider only the residues lining the ligand binding pocket according to the RXR α /9-cis RA complex (22 residues at 4.5 Å shown by green colored dots in Fig. 1), their sequence identity is much higher compared to the whole LBD, around 60 % (12 residues for ctUSP) for dipterans and above 95 % for lepidopterans. In particular, the residues Gln 275 and Arg 316 of hsRXR α anchoring the carboxylate moiety of the 9-cis RA are conserved for all vertebrate and arthropod RXRs except for a few dipteran insects that are discussed later in the text. This observation suggests that most of the USP LBDs could form a similar electrostatic interaction with a ligand.

Comparison of the hvUSP and hsRXR α crystal structures

The crystal structures of hsRXR α and hvUSP are rather similar (Egea et al., 2000; Billas et al., 2001). However, major differences are observed (Figure 2A). The hsRXR α adopts the canonical agonist conformation whereas the hvUSP structure adopts an antagonist conformation in which helix H12 is located in the coactivator groove. The loop 1-3 is displaced in hvUSP compared to hsRXR α and the carboxy-terminal stretch of this loop firmly locks H12 in the antagonist conformation through interactions between residues strictly conserved among lepidopteran and dipteran USPs. The helix H3 is one turn longer than in hsRXR α and its N-terminal region of H3 is shifted outwards the protein core by more than 24 °. It forms together with helices H6 and H11 a large open cavity, which is stabilized by hydrophobic interactions with the phospholipid.

Both in the crystal structures of hvUSP and dmUSP, the fortuitous ligand cocrystallized adopts an identical, but unusual position in the receptor. While in the electron density map of hvUSP, the phospholipid is clearly identified, the situation is more subtle for dmUSP. In fact, six molecules of dmUSP are found in the asymmetric unit of the crystal, i.e. the crystal building block. Each of the six USP units is occupied by a phospholipid molecule that is located at about the same position in the receptor and contacts USP through the same major hydrophobic contacts. The only difference between the six ligand/receptor complexes is the quality of the electron density map observed for the ligand that is mainly influenced by the level of disorder of the lipid inside the receptor. Therefore, for hvUSP and dmUSP, the unusual position of the phospholipid inside the receptor might suggest that the experimental USP structure is stabilized by the phospholipid. The possibility of a structure more closely related to that of RXR should therefore also be examined.

Homology modeling of the USP ligand-binding domains

For the docking studies, three different homology models, i.e. two reference models and one hybrid model have been constructed for several representative members of each major insect order. The first reference model is based on the crystal structure of hsRXR α and the second one on the crystal structure of hvUSP. Finally, a chimeric protein was generated that combines features observed in the crystal structure of hvUSP and in the RXR α homology model of hvUSP. The hybrid molecule is homologous to the hvUSP crystal structure, except that the phospholipid has been

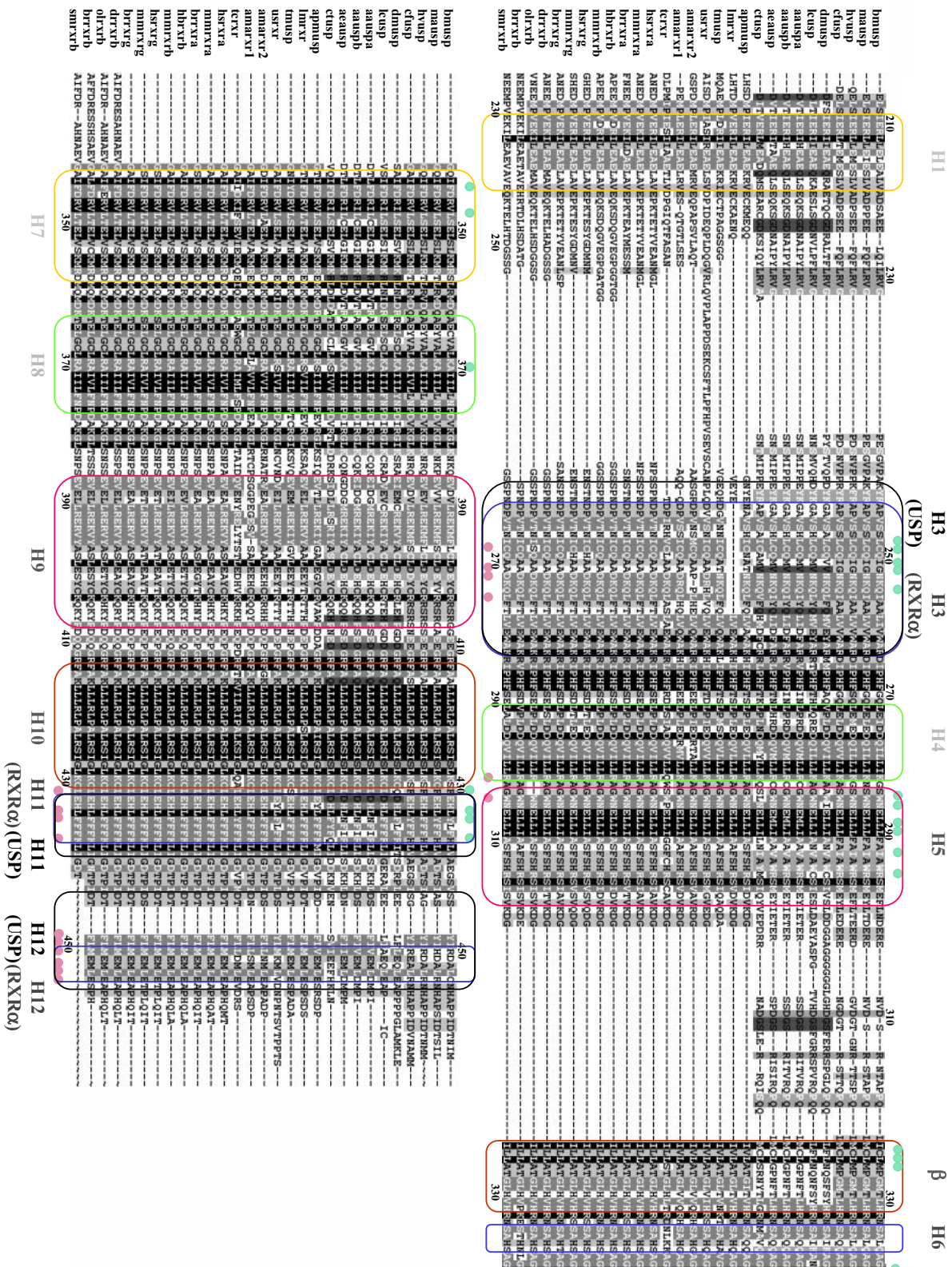


Figure 1. Alignment of USP and RXR LBD sequences. The hvUSP and hsRXR residue numbering are given at the top and bottom of the alignment, respectively. The secondary structure is given as colored boxes according to hvUSP and hsRXRα crystal structures. H indicates α-helices and β the β-sheet. The color code for the amino acid conservation is as follows: white on black : 100 % conservation for USPs-RXR; white on grey: 80 % conservation for USPs-RXR; black on grey: 60 % conservation for USPs-RXR; black on blue: 100 % conservation for Lepidopteran USPs; black on red: 100 % conservation for Dipteran USPs; black on magenta: 100 % conservation for Lepidopteran and Dipteran USPs; white on magenta: 80 % conservation for Lepidopteran and Dipteran USPs. The residues that interact with the 9-cis RA in hsRXRα are indicated by green colored dots, those involved in the stabilization of helix 12 in the agonist conformation are indicated by magenta colored dots. The abbreviations for the different organisms are: bm: *Bombyx mori*; ma: *Manduca Sexta*; hv: *Heliothis virescens*; cf: *Choristoneura fumiferana*; dm: *Drosophila melanogaster*; lc: *Lucilia cuprina*; aa: *Aedes aegypti*; ae: *Aedes albopictus*; ct: *Chironomus tentans*; apm: *Apis mellifera*; lm: *Locusta migratoria*; tm: *Tenebrio molitor*; usrxr: *Uca pugilator*; ama: *Amblyomma americanum*; tc: *Tripedalia cystophora*; hs: *Homo sapiens*; mm: *Mus musculus*; br: *Brachydanio rerio*; dr: *Danio rerio*; ol: *Oryzias latipes*; sm: *Scophthalmus maximus*.

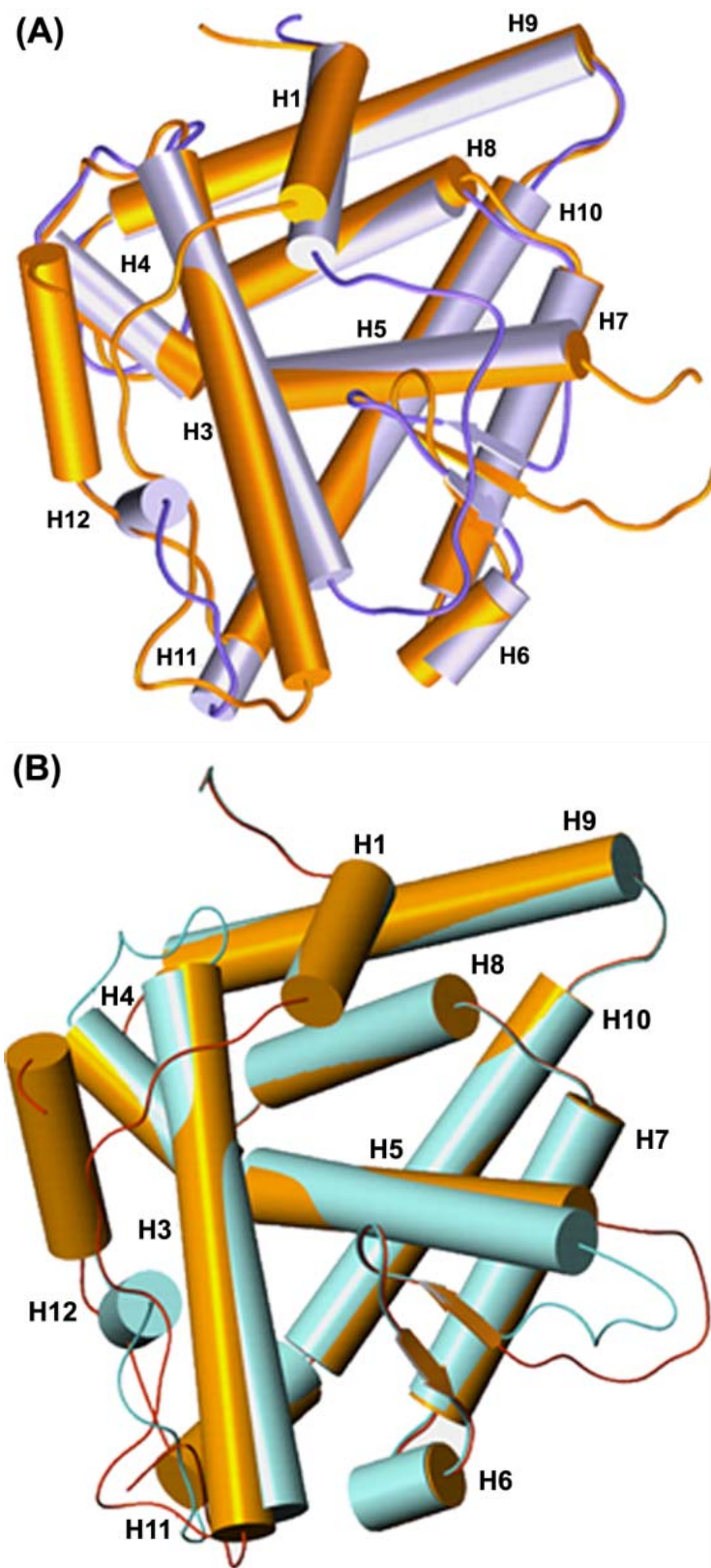


Figure 2. Comparison of the crystal structure of hvUSP LBD with the crystal structure of hsRXR α and with the hvUSP hybrid homology model. α -helices are represented as cylinders and β -sheets as arrows. Figures were generated with Setor.

(A) Superimposition of the LBD crystal structures of hsRXR α in the agonist position (in light purple) and hvUSP in the antagonist position (in orange). (B) Superimposition of the hvUSP LBD crystal structure in the antagonist conformation (in orange) and hvUSP hybrid homology model in the agonist conformation (in light blue).

removed and H12 adopts the agonist conformation (Figure 2B). As a consequence of these two major structural changes, the amino-terminal end of helix H3 is positioned intermediate between the positions observed in the two reference models, and the loop between helices 1 and 3 adopts a rather loose conformation as seen in the hsRXR α structure. The ligand binding pocket of the hybrid model includes, therefore, all of the important features observed in the crystal structure of hvUSP, except for the region left unoccupied after removal of the lipid that is shrunk to the size observed in the RXR structure. This hybrid model was used as a template to generate the hybrid homology models of all the different USPs used in the docking studies.

The helix H12 in the agonist position

Since both crystal structures of hvUSP and dmUSP indicate that H12 adopts an antagonist conformation, it is of interest to understand whether the agonist conformation of H12 in USPs is structurally possible. In order to answer this question, the residues involved in the contacts between H12 in the agonist position and the core of the receptor were examined.

In the agonist conformation of hsRXR α , H12 leans over the ligand binding pocket, and is stabilized by a series of hydrophobic contacts involving residues of H3, H4, H5 and H11 (Egea et al., 2000) (shown by magenta colored dots in Figure 1). Most of these residues are highly conserved or replaced by amino acids of similar nature (Figure 3). For example, adopting the hsRXR α residue numbering, the residues Cys 269 (H3), Trp 305 (H5) and Leu 436 (H11) are strictly conserved in all USPs and RXRs (Figure 1). Furthermore, an aspartic acid in H3 (Asp 273) forms a hydrogen bond between its carboxylate moiety and the backbone NH groups of Phe 450 and Leu 451 just before H12 and stabilizes the agonist position of the activation helix. This residue is strictly conserved in all RXRs, except for tmRXR where it is replaced by an asparagine residue. Similarly, an asparagine residue is found in all USPs (Asn 254 for hvUSP) and it is likely that hydrogen bonds can be formed between the side chain carbonyl group of this residue and the backbone of the loop 11-12. A similar interaction pattern is observed in the estrogen receptor agonist complexes where this aspartate residue has been the focus of numerous mutation studies (Anghel et al., 2000).

Focusing on the amino acid sequence of H12, the motif Glu 453 - Met 454 - Leu 455 - Glu 456 (EMLE) is very well conserved for the RXR sequences, but rather divergent for the USP sequences (EQLE for dmUSP and EEFH for ctUSP). These residues are involved in the interaction of H12 with the protein itself and with cofactors. The Met 454 (H12) is in hydrophobic contact with Trp 305 (H5) and its backbone carbonyl group is hydrogen-bonded through a water molecule to the indole NH group of Trp 305 (H5). The hydrophobic contact is maintained for aaUSP, aeUSP and the lepidopteran USPs. The methionine is replaced by glutamine in dmUSP and lcUSP (EQLE) and by glutamate in ctUSP (EEFH). Examining the three-dimensional model of the receptor indicates that the glutamine and glutamate side chains point towards the surface close to the lysine at the C-terminus of H4 and could interact with its amine group (Figure 3). Clearly, for ctUSP, the carboxylate moiety of the glutamate residue would confer an additional negative charge to the coactivator interaction surface when compared to the

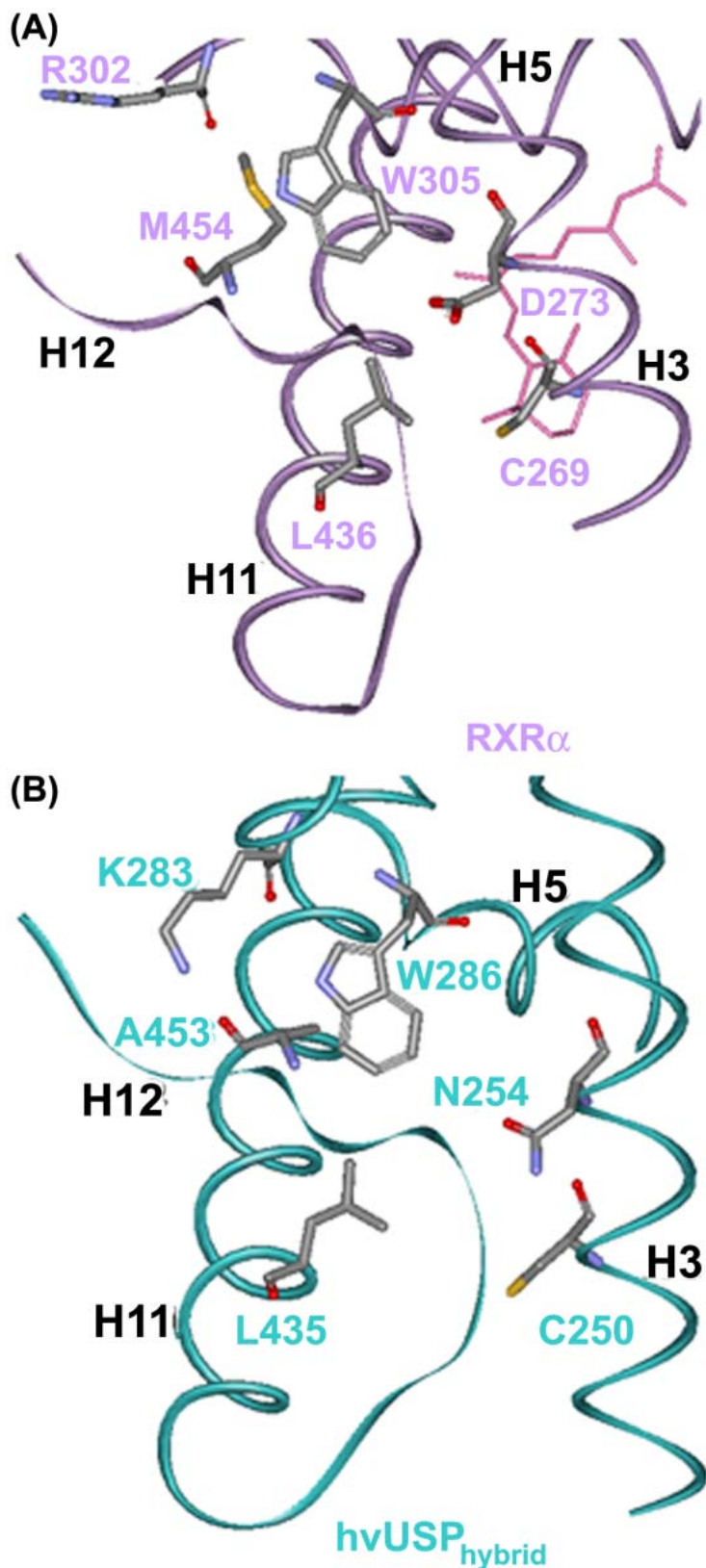


Figure 3. The agonist conformation of the transactivation helix H12 in hsRXR α and in the hvUSP hybrid model. Protein is depicted as a backbone carbon C γ trace. Atoms are represented in standard color scheme: carbon: grey, nitrogen: blue, oxygen: red, sulfur: yellow.

(A) Details of the interactions stabilizing the helix H12 in its agonist position in the RXR α LBD (in light purple). The contacts are due to residues in helices H3, H5, H11 and H12. The side chains of residues participating to the interactions are labeled.

(B) The same region as in (A) for the hvUSP LBD hybrid homology model (in light blue).

other USPs. An additional observation concerns Glu 453 (EMLE in hsRXR α), which together with Lys 284 (H3), have been shown to clamp the LXXLL motif of the coactivator in the groove formed by helices 3, 4, 5 and 12 (Darimont et al., 1998; Feng et al., 1998; Nolte et al., 1998; Shiau et al., 1998). These two residues are conserved or replaced by very similar residues in the various USP sequences (Glu and Arg for dmUSP; Asp and Arg for hvUSP).

In summary, the structural analysis of the residues involved in the interaction of H12 with the receptor suggests that the agonist conformation of USP is possible and that the charge clamp to the coactivator is conserved. However, the coactivator interaction surface is rather different as suggested by the different nature of the residues involved as compared to RXRs. Indeed, in hsRXR α , Glu 281 in H3, Phe 450 in H12 and Met 454 in H12, are respectively replaced by Val 262, Tyr 449, and Ala 453 in hvUSP, and by Glu 294, Leu 490 and Gln 494 in dmUSP.

Putative ligands for USPs

Since no ligands have been clearly identified yet, USP is still an orphan receptor. However, the crystal structures of hvUSP and dmUSP together with the high sequence conservation of the residues lining the LBP suggest that USP is capable of binding ligands. Recent studies suggest that juvenile hormones might be candidate ligands for USP (Jones & Sharp, 1997; Jones et al., 2001). Juvenile hormones are sesquiterpenoids derived from farnesol pyrophosphate. They exhibit a methyl ester group at one end and an epoxide moiety at the other end. Three different forms, JH I, JH II and JH III exist, which differ in the number and location of methyl groups attached to the terpenoid skeleton (see Figure 4A). The juvenile hormones play a role in almost every aspect of insect life. This peculiarity was exploited in the search for insecticides and numerous juvenile hormone mimics were derived, one of which is methoprene (Dhadialla et al., 1998). This ligand, used commercially as an insect growth regulator, closely resembles the juvenile hormones, but lacks the epoxide function (Figure 4A). Other analogs, which differ significantly in structure from juvenile hormones, such as fenoxycarb, pyriproxyfen and diofenolan, also have insecticidal activity. A few of these ligands together with 9-*cis* RA are used for docking studies as described below. Both the ester and the acid forms have been constructed and docked in ligand-binding pocket of the USP homology models.

Ligand binding pocket

The ligand binding pocket in the crystal structure of hvUSP is an open cavity that contains the phospholipid. On the other hand, the cavity is closed in the hybrid homology model and resembles the LBP of the RXR-like homology model, but is slightly wider. Since the residues lining the LBP are well conserved between vertebrate and insect receptors, the size of the cavity is expected to have a similar volume. This is the case for the RXR-like homology models of USP LBDs as shown in Figure 4B. For hsRXR α , the volume of the pocket is 489 Å³, close to the values found for the RXR-like homology models of hvUSP (536 Å³) and other USPs. In the hsRXR α /9-*cis* RA complex, the ligand occupies about 66 % of the cavity (326 Å³). This value lies in the range of typical percentages of LBP occupancy of between 60 and 70 % (63% for ER/estradiol, 67 % for PR/progesterone, 66 % for RAR/all-*trans* RA). Lower

occupancies are found for VDR/vitamin D (56 %) and in the extreme case of PPAR γ /rosiglitazone (40 %). For the USP hybrid homology models, their volume is 10 to 30 % larger compared to RXR-like model (598 Å³ for hvUSP). Consequently, the percentage of occupancy is lower than those observed for the RXR-like models, lying in the lower range of typical values observed for classical nuclear receptors. For example, in the extreme case of JHI, the acidic and ester forms occupy 48 and 54 % of the hvUSP hybrid model LBP, respectively.

Docking of putative ligands

The ligands chosen for the docking study were the juvenile hormones in the acidic and ester forms and the 9-cis RA. The choice of synthetic juvenile hormone analogs was restricted to the acidic and ester form of methoprene. Docking of these ligands was performed for a representative member of lepidopteran USPs (hvUSP). Preliminary results will also be presented for a dipteran USP (dmUSP), but this case still needs further calculations.

As a first step, the RXR-like homology models of USP were chosen for ligand docking. By exploiting the strong conservation of residues involved in the anchoring of the carboxylate

	Gold	Charmm		
ligand	score	E _{inter}	Vdw _{inter}	eel _{inter}
9-cis RA	46.7	-54.8	-37.6	-17.2
JH-I acid	42.1	-53.2	-33.4	-19.8
JH-II acid	42.7	-51.5	-29.1	-22.4
JH-III acid	40.2	-53.5	-31.5	-22.1
methoprene acid	44.5	-63.8	-35.3	-28.5
JH-I ester	43.2	-42.2	-41.4	-0.8
JH-II ester	40.6	-31.7	-35.8	-4.1
JH-III ester	41.8	-39.2	-38.5	-0.7
methoprene ester	30.6	-41.8	-42.0	0.1

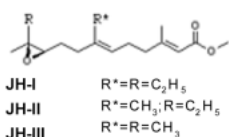
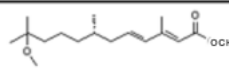
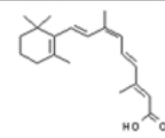
Table 1. Results of the docking of various ligands in the hvUSP hybrid homology model. The score obtained by Gold as well as the total interaction (E_{inter}) energy obtained by Charmm are indicated. E_{inter} comprises the van der Waals energy term (Vdw_{inter}) and the electrostatic energy term (eel_{inter}).

group of the 9-cis RA in RXR α , the acidic forms of the various ligands were expected to bind in a similar fashion to USPs (Figure 5A). However, the arginine residue in H5 is not accessible to the ligand carboxylate moiety in the RXR-like LBP because of the much more constrained region around this anchoring residue. In fact, in RXR α , an alanine (Ala 271 in H3) and a phenylalanine (Phe 313 in H5) are located on both sides of the ligand carboxylate group (Figure 5A). The minimal distance between them is 7.2 Å, which is sufficient for the ligand to be at a van der Waals distance from these two residues. However, these two residues are replaced by two isoleucines for hvUSP (Figure 5B), as is also the case for other lepidopteran USPs, and by two valines for dmUSP (Figure 5C), one of which is conserved and the other one replaced by a methionine in other dipteran USPs. The bulkiness of these residues does not allow the ligand to be accommodated in this restricted cavity, or to interact with the arginine residue. On the other hand, if the RXR-like pocket is compared with that of the hvUSP crystal structure, the latter is wider. In fact, in the hvUSP crystal structure, the isoleucine residues are separated by 7.1 Å and the ligand can fit in this region and interact with the anchoring arginine of H5. These observations give support to the hybrid homology model, which is identical to the hvUSP crystal structure in this region.

All the ligands were docked in these hybrid models using the Gold program. The Gold docking program has its own energy function to evaluate the fitness of the ligand/protein complex. The larger the value, the better the fit. Since Gold considers only the ligand as flexible and the side chains of the protein as rigid, we used the alternative molecular energy function, Charmm, which allows the protein side chains to relax. The results of these two experiments are reported in Table 1. The Charmm results correspond to the interaction energy between the ligand and the receptor (E_{inter}). The best fit was obtained with the largest negative interaction energy. The results of the calculations indicate that the acidic form of the ligands would preferentially bind to hvUSP and most probably also to other USPs that exhibit an arginine residue in H5. In this case, a salt bridge is formed between the negative charge carried by the carboxylate moiety of the ligand and the positive charge carried by the guanidium group of Arg297 (hvUSP) (Figure 5B), as observed for RXR / 9-cis retinoic acid experimental complex (Figure 5A).

On the other hand, the juvenile hormone and methoprene esters cannot form a similar interaction network and the energy penalty amounts to about 15 to 20 kcal/mol, essentially due to the reduced electrostatic contribution. While the acidic functional groups

(A)

Ligands	Formula	Volume (Å ³)	Surface (Å ²)
JH-I JH-II JH-III	 JH-I R* = R = C ₂ H ₅ JH-II R* = CH ₃ ; R = C ₂ H ₅ JH-III R* = R = CH ₃	309 (289) 294 (279) 273 (256)	320 (305) 305 (286) 298 (281)
Methoprene		347 (277)	324 (296)
9-cis RA		326	336

(B)

	RXR α -like	hybrid USP/RXR
RXR α	489	--
hvUSP	536	598
cfUSP	571	580
dmUSP	539	574
aaUSP	555	630
ctUSP	515	587
1cUSP	431	568
amRXR	462	--
tmRXR	455	--

Figure 4. Juvenile hormones (JH) and their analogs.

(A) Chemical structures and sizes of ligands used for docking. The volume and surface of the ligand is reported for the ester form and for the acidic form they are given in parentheses.

(B) Volume of the ligand binding pocket of RXR α and USPs of various insect species, using the RXR α -like models and the hybrid homology models of USPs as discussed in the text.

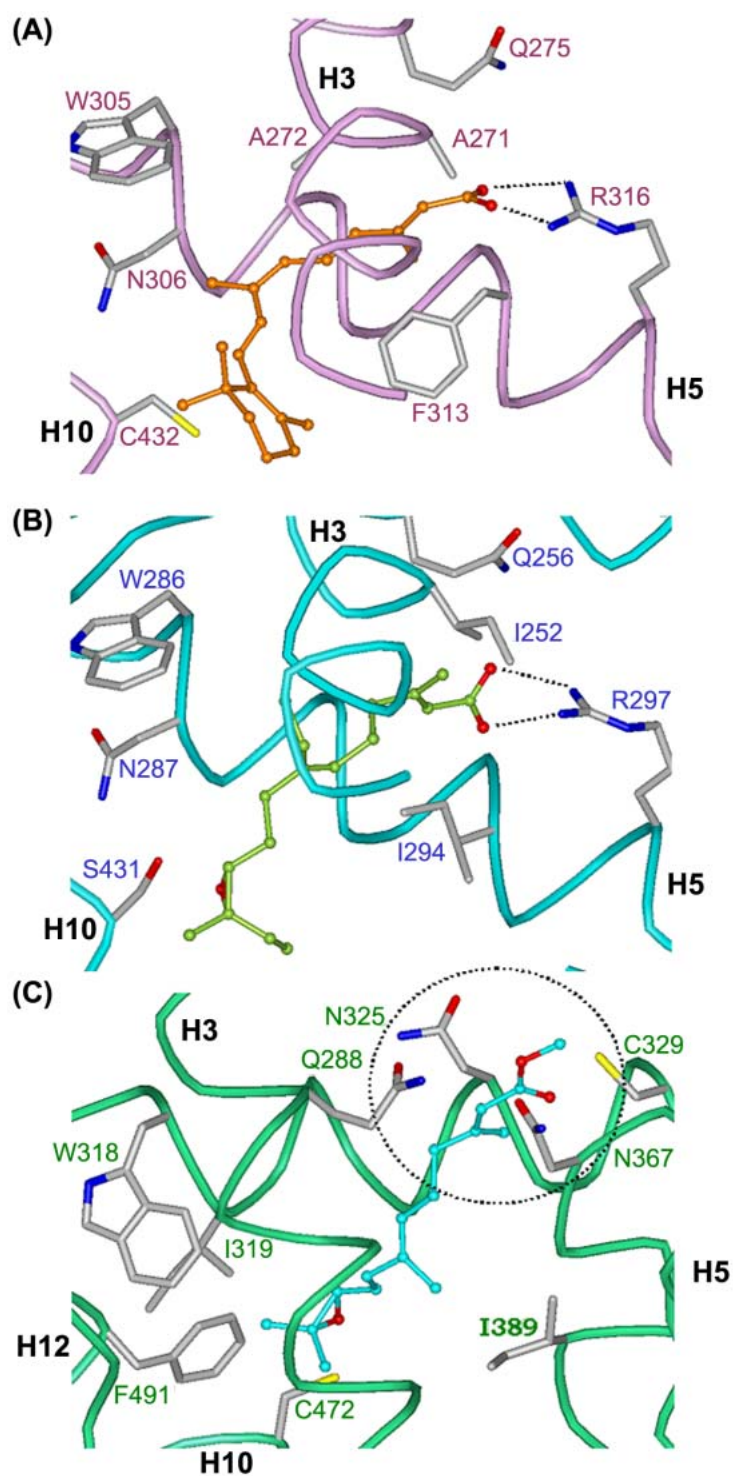


Figure 5. The binding of 9-cis RA to hRXR α and the two JH binding modes in the hvUSP and dmUSP agonist homology models. Protein is depicted as a backbone C α -trace. Helices are indicated. Atoms are represented in standard color scheme: carbon: grey, nitrogen: blue, oxygen: red, sulfur: yellow. Dotted lines indicate H-bonds. The figures have been generated by using WebLab Viewer.

(A) The 9-cis RA in the hRXR α LBP. The protein is colored in light purple. The ligand is depicted in orange. Only a few side chains are labeled for clarity.

(B) Docking of the acidic JH I in the hvUSP LBP agonist model. The protein is colored in light blue. The ligand is colored in green.

(C) Docking of the ester JH III in the dmUSP LBP agonist model. The protein is colored in green. The ligand is colored in cyan. The dotted circle corresponds to the H-bond network region to which the ester group of JH binds.

of the various ligands cluster in the same region, the epoxide moiety of the various juvenile hormones adopt a more scattered distribution, probably due to the rather large cavity in the region around Ser 431 in H10 (hvUSP). The scores evaluated with Gold or with Charmm do not show a clear preference for any of the juvenile hormones, while methoprene acid exhibits the best score.

The type and the nature of the residues lining the LBP of USPs are rather similar, except for the dipteran dmUSP, lcUSP and ctUSP. For these USPs, the arginine in H5, which participates to the anchoring of the ligand carboxylate group is replaced by cysteine or methionine. A careful analysis of the 3D models reveals that the change from a charged (Arg) to an apolar residue (Cys or Met) is correlated to changes in two other residues located in the same region. Firstly, Ala 293 (hvUSP) in H5 is replaced by an asparagine and secondly, Met 326 (hvUSP) in the β -sheet is replaced either by an asparagine for dmUSP and lcUSP or by serine for ctUSP. These changes suggest that the charged interaction network comprising the arginine residue in H5 is replaced by an intricate network of hydrogen bonds with which additional polar residues in H3, H5 and the β -sheet participate. Indeed, the Figure 5C shows that the putative anchoring region in dmUSP is composed of three polar residues: Gln 288, Asn 325 and Asn 367. In such an environment, the negatively charged carboxylate moiety of the ligand cannot be counterbalanced by a positive charge. Hence, it is likely that ligands with an ester functional group would be preferred over ligands bearing a charged carboxylate group. This consideration is supported by recent binding data showing that dmUSP binds JH III ester and not JH III acid, albeit at low affinity (Jones et al., 2001). However, the docking of ligands in these USPs requires a systematic study of the interactions generated by the various rotamers of the residues involved in the ligand binding.

Finally, Table 1 shows that the score of the 9-cis RA in hvUSP is similar to those of juvenile hormones. From the biological point of view, this is unexpected since the 9-cis RA does not seem to have any effects *in vitro* on the transactivation of EcR/USP or on the binding to EcR/USP (Guo et al., 1998; Thomas et al., 1993). In hvUSP, the 9-cis RA interacts through its carboxylate moiety with the arginine of H5, just like in RXR α . However, it adopts a different conformation as that seen in the RXR α crystal structure. In fact, its β -ionone ring points towards H6 and the N-terminal region of H3. In the model, a displacement of the N-terminal part of helix H3 would generate a small cavity more adapted to the size of juvenile hormones and push the epoxide group of these ligands toward the interacting serine residue in H10 (Ser 431 in hvUSP). As a consequence of this rearrangement, the binding of juvenile hormones might be favored while that of the 9-cis RA would be sterically hampered. The construction of such an alternate homology model and the docking of ligands will be the subject of a future study.

Discussion

The crystal structures of hvUSP and dmUSP both show that the helix 12 bearing the ligand-dependent activation function is locked in the antagonist conformation. Furthermore, the alignment of the residues in H12 of USPs and RXRs shows a poor sequence homology. Therefore, it is not obvious how to decipher whether the agonist conformation in USPs is conceivable. This study has

demonstrated that the agonist positioning of H12 in USPs could structurally be possible. However, two direct consequences of this structural rearrangement should be discussed. First, we have seen that the pattern of interactions for the positioning and the stabilization of H12 in the agonist conformation is rather similar to that seen in RXR α . In addition, the charge clamp which positions the LXXLL motif into the coactivator groove of the receptor is conserved in the RXRs. However, the coactivator interaction surface generated by helices 3, 4 and 12 bears residues, which are of different type and nature from those seen in RXR α . Therefore, we might expect that the coactivators that would bind to USPs are different compared to those that bind to RXRs. This is not unexpected since different cofactors are found in insects than in vertebrates (Bai et al., 2000; Beckstead et al., 2001). The second consequence of the agonist conformation of USP concerns the loop 1-3 known to firmly lock H12 in the antagonist position. As suggested by Billas et al. (2001) and Clayton et al. (2001), the high sequence conservation in this region most probably indicates that the residues have acquired a functional and structural role during evolution. Since this loop was modeled after RXR α in the USP hybrid model, the question arises about the role that the loop 1-3 would play in an agonist USP conformation.

USP is most probably a ligand-regulated receptor as suggested by the crystal structures of hvUSP and dmUSP that contain a LBP partly filled by a lipid. The region of the pocket that is similar to that of RXR α is left unoccupied. In order to dock ligands in this unoccupied region, two homology models for USPs have been considered; one based on the RXR α crystal structure and one hybrid model based on the hvUSP crystal structure, but with H12 in the agonist conformation. From the analysis of the residues lining the LBPs of the hvUSP and dmUSP hybrid homology models, two different patterns of ligand binding emerge. For USPs such as hvUSP that encompass a conserved arginine in H5, the ligand binding is most probably similar to that observed for RXRs. It relies on the formation of a salt bridge between the guanidinium group of this residue and a ligand functional group, preferentially a carboxylate group. This is confirmed by the results of juvenile hormones and their analogs docking in the LBP of hvUSP. The acidic forms of the ligands are preferred over the ester forms by about 15 to 20 kcal/mol, corresponding to an affinity constant two to three orders of magnitude larger for the acidic form than for the ester form. The second mode of ligand binding is observed in dmUSP, lcUSP and ctUSP. For these USP members, the arginine in H5 is replaced by an uncharged and non polar residue. A more detailed structure-based analysis underlines correlated changes in other residues located in the same region. Typically, apolar residues are replaced by serine or asparagine, i.e. by polar amino acids. Therefore, the ligand inside the pocket is surrounded by a network of hydrogen bonds generated by these residues. Computationally, the docking of ligands in such an intricate network of interaction is a difficult task. It requires the generation of a very large number of configurations of the residues in the LBP, each being then used as an initial docking configuration.

The docking of ligands in the LBP is an instructive step that can give insight into the nature of putative ligands. Using the crystal structure of RXR α as a template for USP results in a pocket that cannot accommodate ligands of the juvenile hormone type. In particular, the anchoring region at the C-terminus of H5 cannot be

accessed because of steric constraints due to bulky residue side chains not present in RXR α . On the other hand, the hybrid model which is identical to the crystal structure of hvUSP in this specific region is wider and ligands can nicely fit inside the cavity. At the opposite side of the LBP, close to helices 4 and 11, the shape and the volume of the cavity strongly depend on the model template used. In fact, the position of the amino-terminal part of helix 3 modulates the size of the pocket in this region. The rather similar energies of the various acidic juvenile hormones in the LBP of hvUSP do not allow discrimination of the most favored juvenile hormone. In our model, the three types of juvenile hormones are expected to bind to the receptor. It is of interest to compare these predictions with experimental data. For lepidopteran insects, it has been established that JH I and II are predominant, while JH III is present in minor quantity (Truman & Riddiford, 2002). Moreover, in *Manduca sexta*, JH I and II are about 200 times more active than JH III in the black larval assay (Fain & Riddiford, 1975), 300 times more active in the pupal assay (Riddiford & Ajami, 1973), and 500 times more active in the adult assay for egg maturation (Nijhout & Riddiford, 1974). Only JH I and II are found in the *M. sexta* larva, while the adult female corpora allata makes about equal amounts of JH II and III. For dipteran insects, bioassays on *Drosophila melanogaster* pupae show that JH I is also more active than JH III (Postlethwait, 1974). This activity is interesting given that *D. melanogaster* produces mostly JH III and JH III bisepoxide. JH III bisepoxide is nearly 10 times less active than JH III (Richard et al., 1989) and was recently shown to bind dmUSP (Jones et al., 2001). The theoretical docking study of juvenile hormones and their analogs in hvUSP gives results in terms of energy and configurations that suggest that juvenile hormones can fit well inside the LBP of hvUSP. However, the percentage of occupancy of the LBP by these ligands was shown to lie in the bottom range of values for classical nuclear receptors. This observation questions the validity of USP as the juvenile hormone receptor. On the other hand, it is known that orphan receptors such as PPARs have a lower affinity and a poorer specificity to ligands than the more classical endocrine receptors (for example ER, RAR, TR). They are characterized by a large LBP and a low level of occupancy and behave more like sensors than classical high affinity receptors. The low values of LBP occupancy observed for USPs could therefore reflect this tendency. USP is one of the most intriguing and interesting insect nuclear receptors. It plays a fundamental role in insect development and metamorphosis, but its exact role and function has still not been elucidated yet. As suggested by this preliminary study, the juvenile hormones seem to fit nicely inside the ligand binding pocket of lepidopteran USPs. These theoretical findings favor USP as the juvenile hormone receptor. The soundness of this outcome needs to be supported by stringent experimental evidence. It remains true that our results provide a convincing playground for future developments.

References

- Anghel SI, Perly V, Melancon G, Barsalou A, Chagnon S, Rosenauer A, Miller WH, Jr., Mader S. 2000. Aspartate 351 of estrogen receptor alpha is not crucial for the antagonist activity of antiestrogens. *Journal of Biological Chemistry* 275:20867-

- 20872.
- Bai J, Uehara Y, Montell DJ. 2000. Regulation of invasive cell behavior by taiman, a *Drosophila* protein related to AIB1, a steroid receptor coactivator amplified in breast cancer. *Cell* 103:1047-1058.
- Beckstead R, Ortiz JA, Sanchez C, Prokopenko SN, Chambon P, Losson R, Bellen HJ. 2001. Bonus, a *Drosophila* homolog of TIF1 proteins, interacts with nuclear receptors and inhibits betaFTZ-F1-dependent transcription. *Molecular Cell* 7:753-765.
- Billas IML, Moulinier L, Rochel N, Moras D. 2001. Crystal structure of the ligand binding domain of the ultraspiracle protein USP, the ortholog of RXRs in insects. *Journal of Biological Chemistry* 276:7465-7474.
- Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. 1983. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *Journal of Computational Chemistry*. 4:187-217.
- Clayton GM, Peak-Chew SY, Evans RM, Schwabe JWR. 2001. The structure of the ultraspiracle ligand-binding domain reveals a nuclear receptor locked in an inactive conformation. *Proceedings of the National Academy of Sciences USA* 98:1549-1554.
- Darimont BD, Wagner RL, Apriletti JW, Stallcup MR, Kushner PJ, Baxter JD, Fletterick RJ, Yamamoto KR. 1998. Structure and specificity of nuclear receptor-coactivator interactions. *Genes and Development* 12:3343-3356.
- Dhadialla TS, Carlson GR, Le DP. 1998. New insecticides with ecdysteroidal and juvenile hormone activity. *Annual Review of Entomology*. 43:545-569.
- Dunbrack RL Jr., Karplus M. 1993. Backbone-dependent rotamer library for proteins. Application to side-chain prediction. *Journal of Molecular Biology*. 230:543-574.
- Egea PF, Mitschler A, Rochel N, Ruff M, Chambon P, Moras D. 2000. Crystal structure of the human RXRalpha ligand-binding domain bound to its natural ligand: 9-cis retinoic acid. *EMBO Journal* 19:2592-2601.
- Fain MJ, Riddiford LM. 1975. Juvenile hormone titers in the hemolymph during late larval development of the tobacco hornworm, *Manduca sexta* (L.). *Biological Bulletin* 149:506-521.
- Feng W, Ribeiro RC, Wagner RL, Nguyen H, Apriletti JW, Fletterick RJ, Baxter JD, Kushner PJ, West BL. 1998. Hormone-dependent coactivator binding to a hydrophobic cleft on nuclear receptors. *Science* 280:1747-1749.
- Guo X, Xu Q, Harmon MA, Jin X, Laudet V, Mangelsdorf DJ, Palmer MJ. 1998. Isolation of two functional retinoid X receptor subtypes from the Ixodid tick, *Amblyomma americanum* (L.). *Molecular and Cellular Endocrinology* 139:45-60.
- Hendlich M, Lackner P, Weitckus S, Floeckner H, Froschauer R, Gottsbacher K, Casari G, Sippl MJ. 1990. Identification of native protein folds amongst a large number of incorrect models. The calculation of low energy conformations from potentials of mean force. *Journal of Molecular Biology* 216:167-180.
- Hirai M, Shinoda T, Kamimura M, Tomita S, Shiotsuki T. 2002. *Bombyx mori* orphan receptor, BmHR78: cDNA cloning, testis abundant expression and putative dimerization partner for *Bombyx* ultraspiracle. *Molecular and Cellular Endocrinology* 189:201-211.
- Jones G, Sharp PA. 1997. Ultraspiracle: an invertebrate nuclear receptor for juvenile hormones. *Proceedings of the National Academy of Sciences USA* 94:13499-13503.
- Jones G, Willett P, Glen RC, Leach AR, Taylor R. 1997. Development and validation of a genetic algorithm for flexible docking. *Journal of Molecular Biology* 267:727-748.
- Jones G, Wozniak M, Chu Y, Dhar S, Jones D. 2001. Juvenile hormone III-dependent conformational changes of the nuclear receptor ultraspiracle. *Insect Biochemistry and Molecular Biology* 32:33-49.
- Jones TA, Zou JY, Cowan SW, Kjeldgaard M. 1991. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta crystallographica A* 47:110-119.
- Laskowski RA, MacArthur MW, Moss DS, Thornton JM. 1993. PROCHECK: A program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography* 26:283-291.
- McCarty MF. 2001. The chlorophyll metabolite phytanic acid is a natural rexinoid—potential for treatment and prevention of diabetes. *Medical Hypotheses* 56:217-219.
- Nicholls A, Sharp KA, Honig B. 1991. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* 11:281-296.
- Nijhout MM, Riddiford LM. 1974. The control of egg maturation by juvenile hormone in the tobacco hornworm moth, *Manduca sexta*. *Biological Bulletin* 146:377-392.
- Nolte RT, Wisely GB, Westin S, Cobb JE, Lambert MH, Kurokawa R, Rosenfeld MG, Willson TM, Glass CK, Milburn MV. 1998. Ligand binding and co-activator assembly of the peroxisome proliferator-activated receptor-gamma. *Nature* 395:137-143.
- Oro AE, McKeown M, Evans RM. 1990. Relationship between the product of the *Drosophila* ultraspiracle locus and the vertebrate retinoid X receptor. *Nature* 347:298-301.
- Postlethwait JH. 1974. Juvenile hormone and the adult development of *Drosophila*. *Biological Bulletin* 147:119-135.
- Richard DS, Applebaum SW, Sliter TJ, Baker FC, Schooley DA, Reuter CC, Henrich VC, Gilbert LI. 1989. Juvenile hormone bisepoxide biosynthesis in vitro by the ring gland of *Drosophila melanogaster*: a putative juvenile hormone in the higher Diptera. *Proceedings of the National Academy of Science U. S. A* 86:1421-1425.
- Riddiford LM, Ajami AM. 1973. Juvenile hormone: its assay and effects on pupae of *Manduca sexta*. *Journal of Insect Physiology* 19:749-762.
- Riddiford LM, Cherbas P, Truman JW. 2001. Ecdysone receptors and their biological actions. *Vitamins and Hormones* 60:1-73.
- Sali A, Blundell TL. 1993. Comparative protein modelling by

- satisfaction of spatial restraints. *Journal of Molecular Biology* 234:779-815.
- Shiau AK, Barstad D, Loria PM, Cheng L, Kushner PJ, Agard DA, Greene GL. 1998. The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen. *Cell* 95:927-937.
- Sutherland JD, Kozlova T, Tzertzinis G, Kafatos FC. 1995. *Drosophila* hormone receptor 38: a second partner for *Drosophila* USP suggests an unexpected role for nuclear receptors of the nerve growth factor-induced B-type. *Proceedings of the National Academy of Science USA* 92:7966-7970.
- Thomas HE, Stunnenberg HG, Stewart AF. 1993. Heterodimerization of the *Drosophila* ecdysone receptor with retinoid X receptor and ultraspiracle. *Nature* 362:471-475.
- Truman JW, Riddiford LM. 2002. Endocrine insights into the evolution of metamorphosis in insects. *Annual Review of Entomology* 47:467-500.
- Wurtz JM, Bourguet W, Renaud JP, Vivat V, Chambon P, Moras D, Gronemeyer H. 1996. A canonical structure for the ligand-binding domain of nuclear receptors. *Nature Structural Biology* 3:87-94.
- Yao T-P, Forman BM, Jiang Z, Cherbas L, Chen J-D, McKeown M, Cherbas P, Evans RM. 1993. Functional ecdysone receptor is the product of EcR and ultraspiracle genes. *Nature* 366:476-479.