CRITICAL ROLE OF DESOLVATION IN THE BINDING OF 20-HYDROXYECDYSONE TO THE ECDYSONE RECEPTOR

Christopher Browning1*, Elyette Martin1,2*, Caroline Loch1†, Jean-Marie Wurtz1, Dino Moras1, Roland H. Stote2, Annick P. Dejaegere1 and Isabelle M.L. Billas1

From IGBMC1 (Institut de Génétique et de Biologie Moléculaire et Cellulaire), Département de Biologie et de Génomique Structurales, Illkirch, F-67400 France; INSERM, U596, Illkirch, F-67400 France; CNRS, UMR7104, Illkirch, F-67400 France; Université Louis Pasteur, Faculté des Sciences de la Vie, Strasbourg, F-67000 France, Laboratoire de Biophysicochimie Moléculaire2, Institut de Chimie de Strasbourg, Université Louis Pasteur, Strasbourg, France.

†Present address: ZoBio BV, Einsteinweg 55, 2333 CC Leiden, The Netherlands.

*These authors contributed equally to this work; C.B. for crystallography; E.M. for theoretical studies.

Address correspondence to: Dino Moras, Département de Biologie et de Génomique Structurales, IGBMC, 1, rue Laurent Fries, F-67400 Illkirch Cedex, France ; Tel. : 33-3-88-65-32-20; Fax : 33-3-88-32-76 ; E-mail : moras@igbmc.u-strasbg.fr and Annick P. Dejaegere, Département de Biologie et de Génomique Structurales, IGBMC, École Supérieure de Biotechnologie de Strasbourg (ESBS), Parc d’Innovation, Boulevard Sébastien Brant, BP 10413 - F-67412 Illkirch Cedex, France ; Tel. : 33-3-90-24-47-21 ; Fax: 33-3-90-24-47-18; E-mail : annick@igbmc.u-strasbg.fr

Keywords: Ecdysone receptor; EcR; 20-hydroxyecdysone; nuclear receptors; ecdysteroid; hormone binding; signal transduction; desolvation effects; free energy decomposition methods; docking methods; binding affinity; 1,2 diol;

The insect steroid hormone 20-hydroxyecdysone (20E) binds to its cognate nuclear receptor composed of the ecdysone receptor (EcR) and Ultraspiracle (USP) and triggers the main developmental transitions, in particular molting and metamorphosis. We present the crystal structure of the ligand-binding domains of EcR/USP in complex with 20E at 2.4 Å resolution and compare it to published structures of EcR/USP bound to ponasterone A (ponA). PonA is essentially identical to 20E, but lacks the 25-OH group of 20E. The structure of 20E-bound EcR indicates that an additional hydrogen bond is formed compared to the ponA-bound receptor, yet, paradoxically, ponA has a significantly higher affinity for EcR than 20E. Theoretical studies based on docking and free energy methods lead to a rationale for understanding the difference in binding affinities between 20E and ponA. Results of the calculations indicate that the favorable contribution from the extra H-bond made by 25-OH of 20E is counterbalanced by its larger desolvation cost compared to that of ponA. The contribution of 25-OH to the binding affinity is further compared to those of 20- and 22-OH groups. Ligands that lack the 20- or 22-OH group are indeed known to bind less favorably to EcR than 20E, an effect opposite to that observed for ponA. The results indicate that their respective contributions to receptor-ligand complex stability reside mostly in their different contributions to solvation/desolvation. Together, the data demonstrate the critical role of ligand desolvation in determining binding affinity, with general implications for the binding of hormones to their cognate nuclear receptors.

Ecdysteroids are the steroidal molting hormones of arthropods that encompass organisms as diverse as insects and crustaceans. These steroid hormones are key regulators of the major developmental transitions, notably molting and metamorphosis as well as cell differentiation and reproduction. All these events are initiated by the secretion of a brain neuropeptide, the prothoracicotropic hormone that acts on the ring gland, causing it to synthesize and release the steroid hormone ecdysone (E) (1). E is then converted by the cytochrome P450 enzyme ecdysone-20-monooxygenase in the fat body and in peripheral tissues to the biologically active metabolite 20-hydroxyecdysone (20E) (2). The molecular basis of the action of 20E relies upon the binding of this hormone to its cognate nuclear hormone receptor, composed of the ecdysone receptor (EcR) and of Ultraspiracle (USP), the homolog of the vertebrate retinoid X receptor (RXR). In fact, it has been shown that USP is an obligatory heterodimeric partner of EcR, required for both high-affinity ligand and DNA binding (3-6). Upon ligand binding, the heterodimeric nuclear receptor complex EcR/USP triggers cascades of gene expression leading to the major insect developmental transitions. The paucity in the types of hormones responsible for insect development and homeostasis contrasts with the diversity of vertebrate hormones that control a large variety of gene-regulatory pathways. The functional diversity of the ecdysone system is likely due to different and possibly multiple mechanisms. For example, the various isoforms of EcR and USP were shown to be expressed differentially in different tissues and at
biologically, ponA has a much higher binding affinity for EcR than 20E (12). To address this paradox, the crystal structures of the EcR/USP bound to 20E and ponA were used in docking and free energy calculations. The results of the calculations lead to a rationale for interpreting these apparently contradictory observations. These results suggest that, to a significant extent, ligand desolvation effects account for the difference in binding affinities between 20E and ponA. Furthermore, a quantitative analysis of protein-ligand interactions is provided and the contributions of various ligand functional groups to the interaction with the receptor are analyzed. Results of the calculations further emphasize the importance of desolvation energy in the binding of hormones to their cognate nuclear receptor.

**Experimental Procedures**

*Protein expression and purification* - The *Heliothis virescens* USP-LBD (residues 205-466) was expressed from the T7 promoter of plasmid vector pACYC11b with chloramphenicol resistance. The wild type HvEcR-LBD (residues 284-532) was expressed as a hexahistidine tagged protein fused to thioredoxin from the expression vector pET32b (Novagen) that conferred ampicillin resistance. *Escherichia coli* strain BL21(DE3) cells (Novagen) transformed with the USP-LBD and EcR-LBD expression plasmids were grown in 2xLB medium at 37°C and subsequently induced for 18 hours with 0.33 mM isopropyl-β-D-thiogalactopyranoside at 18°C. Cell pellets were resuspended in binding buffer (20 mM TRIS [pH=8.0], 50 mM NaCl, 150 mM KCl, 4 mM CHAPS, 2 % glycerol) and clarified by sonication.

**Crystallization** - The purified heterodimer complex was eluted using 125 mM imidazole (pH=8.0). After tag removal by thrombin digestion, the protein was subjected to gel filtration on a Superdex 200 16/60 column (Pharmacia Biotech). The 1:1 molar ratio of EcR-LBD and USP-LBD was confirmed by N-terminal amino-acid sequencing, SDS-PAGE analysis and electrospray ionization mass spectrometry. The heterodimer complex was eluted using 125 mM imidazole (pH=8.0). After tag removal by thrombin digestion, the protein was subjected to gel filtration on a Superdex 200 16/60 column (Pharmacia Biotech). The 1:1 molar ratio of EcR-LBD and USP-LBD was confirmed by N-terminal amino-acid sequencing, SDS-PAGE analysis and electrospray ionization mass spectrometry.

**Crystallization** - The purified heterodimer complex was conditioned in 20 mM TRIS (pH=8.0), 50 mM NaCl, 150 mM KCl, 4 mM CHAPS, 2 % glycerol (v/v) and concentrated to 4-6 mg/ml. Crystallization experiments were carried out at 24°C using both the hanging-drop vapor and the sitting-drop diffusion methods. The crystallization conditions were 30 % PEG 4000, 0.2 M ammonium sulfate, 0.1 M Na citrate (pH=5.6). Crystals were flash-frozen in liquid nitrogen after a short dip in a solution containing 15 % ethylene glycol and 5 % PEG 400 as a cryoprotectant.

**Data collection, structure determination and refinement** - Two sets of data were used for structure refinement.
refinement, a first one at a resolution of 2.85 Å on ID14-EH3 at ESRF (Grenoble, France) and subsequently, a second data set at a higher resolution of 2.4 Å was collected at the Swiss Light Source (Villigen, Switzerland). Data were treated using HKL 2000 (16). The crystal structure was solved by molecular replacement using Amore (17) with the ponA-bound HvEcR/HvUSP crystal structure as a search model. A clear solution with one heterodimer in the asymmetric unit was obtained in the space group P3_1 21 with a correlation coefficient of 65.0 % and a R-factor of 32.5%. Rigid body refinement resulted in R and Rfree values of 29.5% and 31.1%, respectively. A refined structure corresponding to the 2.85 Å dataset was used as a model to carry out molecular replacement using Molrep (18) for the dataset at the higher 2.4 Å resolution. This resulted in a solution having a correlation coefficient of 74.7% and R-factor of 35.3%. Refinement was performed with CNS using torsion angle dynamics with a maximum likelihood function target, including restrained anisotropic refinement of individual atomic temperature factors and bulk solvent correction (19). This resulted in a structure with good R and Rfree values of 24.4% and 30.1%, respectively. Final refinement was performed using REFMAC5 (20) included in the graphical CCP4 interface and the ligand library generated by PRODRG (21) resulting in final R and Rfree values of 22.4% and 28.2%, respectively. Manual adjustments and rebuilding of the models were performed using the program O (22) and sigmaA-weighted electron density maps. The final model comprises for EcR residues maps. The final model comprises for EcR residues Val287-Thr322, Asp333-Ala529 and for USP residues Gln206-Thr303, Thr316-Arg455. The missing regions corresponded to flexible loops between H2 and H3 for EcR and between H5 and the β-sheet for USP. Structures exhibit good geometry with no Ramachandran outliers. Molecular graphics figures were created using the PyMOL Molecular Graphics System (DeLano Scientific, San Carlos, CA, USA; http://www.pymol.org).

Chemicals - 20-Hydroxyecdysone was purchased from SciTech (Prague, Czech Republic) and dissolved in ethanol.

Ligand / protein interactions using scoring functions - The ligands ponA and 20E were docked to their respective crystal structures using the docking programs GOLD (23) and SURFLEX (24,25), which have been recently reported to give the most accurate results for ligand conformational searching and ligand-protein complex energy scoring (26). The respective scoring functions of the two algorithms were used to evaluate the docked structures. Included in the scoring functions are terms for hydrophobic complementarity, polar complementarity, entropic terms, and solvation terms. The respective GOLD and SURFLEX scoring functions were used to rank multiple docked ligand conformations that were within a 1 Å root mean-square deviation (r.m.s.d.) of the experimental ligand position in the binding pocket. Default values for all parameters and a cavity centered on the experimental position of the ligand were used in the docking experiments. For each given scoring function, the total scores, as well as the individual components of the interaction energies did not vary significantly among closely related conformations and therefore, only one conformation will be discussed in what follows.

Free energy decomposition of interactions between HvEcR LBD and 20E or ponA - To obtain a semi-quantitative estimate of the contributions of the various amino acids in the ligand binding pocket to the binding stabilization of 20E and ponA, an MM/PBSA analysis was done following the approach presented by (27). Briefly, we present the procedure followed specifically for this study. Starting from the crystal structure of the HvEcR LBD in complex with 20E, hydrogen atoms were added using the HBUILD (28) module of the CHARMM (Chemistry at HARvard Macromolecular Mechanics) program (29). The results were visualized using PyMOL Molecular Graphics System (DeLano Scientific, San Carlos, CA, USA; http://www.pymol.org) to ensure proper positioning of important hydrogen bond interactions. The hydrogen atom on the hydroxyl group at position C25 (see scheme in Fig. 2) of 20E forms an H-bond with the oxygen atom of the side chain amide group of Asn 504; the geometry of this H bond is: length H-O = 1.62 Å and angle O-H-O = 165°. A similar procedure was done for HvEcR LBD in complex with ponA, e.g., identical initial protein structures were used for both complexes. After addition of the hydrogen atoms, the crystal structures were energy minimized according to the following protocol: harmonic constraints were placed on the protein heavy atoms (50 kcal/mol·Å² for backbone and 25 kcal/mol·Å² for side chains at the beginning). The steroid rings of the ligand ponA and 20E were kept fixed and the system was energy-minimized with 1000 steps of Steepest Descent minimization followed by 1000 steps of Adopted Basis Newton-Raphson minimization. Every 200 steps the harmonic constraints were decreased by 35 %. This energy minimized structure was used in the MM/PBSA analysis, for which the electrostatic component was calculated using the continuum electrostatics program UHBD (30). Final focusing grids of 0.4 Å and 0.34 Å were used for 20E and ponA, respectively, with a probe radius of 1.4 Å. The dielectric constant used for the solvent was 80 and for the solute (protein and ligand), values of 1, 2, 4 and 8 were used. Multiple values were used to ensure robustness of the results. Crystallographic water molecules with low B factors present in the LBD binding pocket were included in the calculations as part of the protein structure, and given a low dielectric constant. The atomic charges and radii used for the protein were those of the CHARMM 22 force field (31), for water the CHARMM modified TIP3 model (32) was used. Electrostatic interactions were decomposed into
 RESULTS

Crystal structure of HvEcR/HvUSP in complex with 20E: Heterodimer structure. The ligand binding domains of HvEcR and HvUSP (residues 284-532 for HvEcR and 205-466 for HvUSP) were co-expressed in Escherichia coli, purified in the presence of the EcR agonist 20E, and crystallized in the P3_121 space group. Crystal diffracted up to 2.4 Å resolution, and the full data set was considered for structure refinement. The heterodimer structure was solved by molecular replacement using a model based on the X-ray structure of HvEcR/HvUSP bound to ponA (13). A molecular replacement solution with a high correlation coefficient was readily found and encompassed one molecule in the asymmetric unit of the P3_121 space group. The data and refinement statistics are summarized in Table 1.

The overall structure of the heterodimer HvEcR/HvUSP bound to 20E is highly homologous to the structure of the ponA bound HvEcR/HvUSP complex (Figure 1A). When viewed in an orientation that is similar to the other nuclear receptor heterodimer structures (38-42), it is butterfly-shaped. The structure of Heliothis USP in the complex is almost identical to the structure of the monomer (rms deviation of 0.71 Å for 235 Cα atoms) (43), with a large hydrophobic ligand binding pocket that is filled with a phospholipid molecule. A similar X-ray structure was also observed for Drosophila USP LBD, while a different structure was observed for Bemisia USP (14) and Tribolium USP (15).

The structure of HvEcR bound to 20E is highly similar to the structures of Heliothis EcR, Bemisia EcR and Tribolium EcR bound to ponA (rms deviation of 0.5 Å for 233 Cα atoms) (44). For these EcR LBD structures, the β-sheet is composed of three strands. This is a key region of the receptor that displays a remarkable flexibility in response to different ligand types. In fact, for synthetic DBH agonists, a complete rearrangement of the region of the β-sheet was observed, involving two aromatic residues F397 and Y403 of the β-sheet that changed conformation in order to fill the space left unoccupied by the small DBH molecule (13). Another critical region is the connection between helices H1 to H3, which in 20E and in ponA bound EcR structures features a small helix H2 that is absent in the complex with the DBH compound. The existence of this small helix seems to be intimately linked to the existence of direct and indirect stabilizing interactions with the ligand. Since the small DBH molecule is not located in close vicinity to the region comprising helix H2 and the β-sheet, H2 lacks ligand-mediated stabilizing interactions and unfolds into a loop (44). In the final stages of structure refinement, additional electron
densities were observed in several regions on the solvent-exposed surface of the receptor (Figure 1B). Firstly, a glycerol and a water molecule were identified in the region corresponding to the hydrophobic co-activator binding cleft. Secondly, a well-defined density observed between helices H1 and H9 was attributed to a $SO_4^{2-}$ ion originating from the ammonium sulfate containing crystallization solutions. This sulfate anion makes electrostatic interactions with the amino group of Asn 293 (H1) and with the side chains of Thr343 and Thr346 (H7). Finally, a rather large branched-shaped electron density was seen between the loop H3-H4 and helix H9. We attributed this density to the presence of a citric acid molecule, since crystals were grown in 0.1 M sodium citrate buffer pH=5.6. Remarkably, this molecule is in van der Waals contacts only with EcR and makes no direct H-bond interactions with residues of either the protein or the symmetrical EcR/USP heterodimer molecules.

**Crystal structure of HvEcR/HvUSP in complex with 20E: Ligand binding pocket.** The ligand-binding pocket of HvEcR bound to 20E is rather similar in size and shape to that of the ponA-bound HvEcR (601 Å$^3$). In addition, the interactions observed between ponA and the protein, and in particular the hydrogen bonds, are preserved in the 20E bound complex. However, since the resolution of the X-ray structure of the 20E bound EcR/USP complex is higher than that of the ponA-bound complex (2.4 Å rather than 2.9 Å), a more accurate analysis of the protein/ligand interaction distances is possible (Fig. 2). The three hydroxyl groups of the steroid core (2-β, 3-β and 14-αOH) are hydrogen-bonded to EcR residues. The 2-βOH and the 3-βOH are hydrogen-bonded to the guanidinium moiety of Arg383 (2.8 Å), the 3-βOH to the backbone carbonyl group of Glu309 (2.6 Å) and the 14-αOH to the side chains of Thr343 and Thr346 (3.1 and 3.3 Å, respectively). In addition, the C6-ketone moiety is H-bonded to the amide group of Ala398 (2.7 Å). The alkyl tail of 20E features three hydroxyl groups (20-, 22- and 25-OH), one more (25-OH) than in ponA. The 20-OH group is H-bonded to the side chain of Tyr408 (2.8 Å), as also observed in the ponA-bound EcR structure. Relative to the ponA complex, an additional H-bond is formed between the 25-hydroxy group of 20E and Asn504 having an interaction distance of 2.6 Å, as shown in Figure 3. The 25-OH group is also in van der Waals contact with the C3 of W526 (distance 3.0 Å), in a geometry similar to that observed for water-aromatic interactions (45). Furthermore, this aromatic residue, which was shown to be crucial for the transactivation capability of EcR (46), makes electrostatic interactions with the side chain amino group of Asn504 (H10), while its imino group moiety is H-bonded to the OH group of Ser376 (H5), which itself is H-bonded to the main chain carbonyl group of Leu372 (H4) (Figure 3). Concerning the 22-hydroxyl moiety, it is not directly H-bonded to the protein, in contrast to the 20- and 25-OH groups, but makes water-mediated interactions with EcR amino acid residues. This network comprises 22-OH, the side-chain carbonyl group of Asn504 and the main chain carbonyl groups of Leu500 (in H10) and Val416 (in H7) (Figure 3). No water molecules were reported in the structures of *Heliothis* and *Bemisia* EcR/USP bound to ponA. However, reexamination of the published ponA-bound *Heliothis* EcR/USP structure in the light of the 20E bound structure reveals a small peak in the difference electron density map close to Asn504, suggesting the possible existence of a water molecule at this location. In the case of ponA bound *Tribolium* EcR/USP solved at 2.7 Å resolution, several water molecules were observed in the electron density that form a tight interaction network bridging together the 22-OH group with EcR residues of H7 and H10-H11 in a way similar to that observed for the 20E-bound EcR structure.

**Crystal structure of HvEcR/HvUSP in complex with 20E: Ligand conformation.** Inside the pocket, 20E is oriented with its steroid A-ring towards H1 and the β-sheet, while the alkyl chain points towards the C-terminal helix H12. The steroid core of 20E adopts a chair conformation with the 2-βOH and the 3-βOH groups in equatorial and axial orientations, respectively. This steroid core conformation is similar to that observed in the crystallographic structure of the free 20E molecule (47). The free 20E molecule was crystallized under varying hydration conditions, from a non-hydrated state to mono- and tri-hydrated states. Interestingly, the conformation observed for bound 20E is similar to those observed in the hydrated 20E crystals. While in the non-hydrated 20E crystal structure, a network of hydrogen bonds link together the 20-, 22- and 25-OH groups, a rearrangement of the alkyl side-chain conformation occurs in the free hydrated 20E structures (Figure 4). In particular, a re-positioning of the 25-OH takes place allowing this OH-group to interact with a tight network of water molecules and hydroxyl groups belonging to symmetrical molecules. As a consequence, the H-bond between the 22- and 25-OH groups that is observed for the non-hydrated 20E is not present in the hydrated 20E crystal structures (distance O22-O25 of 4.64 Å, instead of 2.67 Å for the non-hydrated crystal). Thus, the crystallographic data of the free 20E molecule suggests the importance of water molecules in the proximity of 25-OH, while the conformation of the 20- and 22-OH are not very sensitive to their solvation state. Most importantly, in the 20E-bound EcR structure, the conformation of the alkyl side-chain of 20E is very similar to that observed in the hydrated free 20E crystals with a similar O22-O25 distance (4.8 Å).

The structural analysis of the 20E-bound EcR, therefore, indicates that the binding mode of 20E to EcR is almost identical to the binding mode of ponA to EcR, except for the existence of an additional hydrogen bond between the 25-OH group.
of 20E and Asn504. Therefore, a simple rational where the binding affinity is proportional to the number of hydrogen bonds cannot account for the biological data indicating that ponA has a much higher affinity than 20E by one to two orders of magnitude. On the other hand, the comparison with free 20E crystal structures suggests that (de)solvation effects might play a crucial role. In order to examine these effects, we employed theoretical calculations based on docking and free energy methods; the results are presented below.

**Scoring 20E and ponA inside the EcR Ligand Binding Pocket.** The ligand-protein interaction energy was first evaluated using the energy scoring functions of the two docking programs GOLD (23) and Surflex (24,25). The scoring functions of these two programs lend themselves well to the current protein-ligand system given that the nature of the ligand binding pocket is mostly hydrophobic with a few hydrogen bonding sites. This is of particular relevance for the Surflex docking given that this approach has been shown to be appropriate for situations where hydrophobic interactions are predominant. On the one hand, the additional 25-OH group of 20E is exploited by the GOLD program during the docking procedure. The ligands 20E and ponA were docked to their respective crystal structures using these two programs. In both cases, 20E was ranked the better ligand over ponA, with GOLD and Surflex scores of -82.9 [kcal/mol] and 13.8 [-log Kd], respectively for 20E versus scores of -77.3 [kcal/mol] and 12.5 [-log Kd], respectively for ponA. These results are, interestingly, in contrast to the experimental studies, which, as discussed above, demonstrated that ponA has a higher binding affinity toward EcR than 20E.

While in contrast to the experimental observation for binding affinities of 20E and ponA, the docking results do underline the important contribution of the additional 20E hydrogen bond to the binding affinity. The decomposition of the GOLD scoring function into van der Waals, internal torsion and H-bond contributions indicated that the better score obtained for 20E was related to the contribution of the extra hydrogen bond formed between its 25-OH group and Asn504 (5 kcal/mol in favor to 20E; see table S6 in Supplementary Material), while the Surflex steric term indicated no steric clash between the ligand and protein. The non-covalent score, which models hydrophobic complementarity, polar complementarity, entropic terms, and solvation terms (24), also favored 20E over PonaA, albeit by a lesser amount (1 kcal/mol) compared to the GOLD score (5 kcal/mol). This reversed order of affinity observed with both algorithms suggests that effects that are neglected or poorly treated with these models, in particular, (de)solvation, may indeed be important and must be taken into account. Therefore, to address the conundrum of ponA/20E binding affinity, more detailed free energy-based methods were then employed. These results are described in the next section.

**Computational Analysis of the Interactions between 20E and the Ligand Binding Pocket.** The interactions between 20E and the LBP of EcR were evaluated using the free energy decomposition approach described in the Materials and Methods. In the present section, the van der Waals and electrostatic interactions between the ligand 20E and the protein are discussed; desolvation effects are discussed in the following section.

The van der Waals and electrostatic contributions to the protein-ligand interaction are given in Fig. 5. Dominant van der Waals interactions are made by Ile339, Met342, Met380, Val384 and Arg387. Visualisation of the structure shows that the principal interactions made by these respective protein residues are to the steroid skeleton; interaction with the alkyl tail occurs to a much lesser extent. The residues Thr343 and Thr346, which are H-bonded to 14-αOH of 20E and Phe397, which interacts with the steroid A- and B-rings, also make significant contributions to van der Waals interactions. The main favorable electrostatic contributions to the ligand-protein interaction energy are made by residues that are involved in the hydrogen bond network with the ligand. In particular, the steroid core makes strong electrostatic interactions with Thr346 and Arg383 from EcR. One exception is Thr343 that together with Thr346 is H-bonded to 14-αOH, but in marked contrast to it contributes rather unfavorably to the electrostatic interaction energy. Interestingly, a non negligible contribution to the electrostatic interaction energy is observed for Arg387, which is not directly H-bonded to 20E, but rather makes long range stabilizing interactions with the 2-β and 3-βOH groups of the steroid core. Strong electrostatic interactions of the alkyl tail are made with Tyr408, which interacts with the 20-OH group, and with Asn504, which interacts with the 25-OH group.

Overall, the total van der Waals and electrostatic interaction energy terms are about -60 and -40 kcal/mol, respectively, and thus of comparable magnitude. These numbers are indicative of levels of interaction as they are computed on a static crystal structure and reflect only some of the contributions to the total binding free energy. Nevertheless, they attach a quantitative evaluation of the interaction energy to the observation made based on the crystal structures.

**Comparison of Binding Affinities of 20E and ponA: Desolvation effects.** The relative binding affinities of 20E and Pona to the LBP of EcR were compared based on the results of the free energy decomposition calculation. We considered 20E and ponA, both free and bound inside the ligand binding pocket. In interpreting the results, we focused on the electrostatic contributions to their respective binding free energies, and in particular, the compensation...
between ligand-protein interactions and desolvation. Given the similarity between the two ligands, with the exception of the extra hydroxyl group of 20E, it can indeed be assumed that other contributions to the binding free energy will be similar for the two molecules and they will not make a major contribution to the relative binding affinity. The desolvation costs for both ligands were estimated using the UHBD program, the results are given in Table 2. The choice of the internal dielectric constant for the solute has an important influence on continuum electrostatics calculations (see Materials and Methods) and values of 1, 2, 4, and 8 were tested. While comparison with experimental data indicated that a value of 2 was most appropriate for obtaining absolute values of the solvation energies, relative values of solvation free energies are less sensitive to the choice of $\varepsilon$. Using $\varepsilon$ values of 1, 4 or 8 did not change the general results of the calculations, given that only relative values of the desolvation and interaction free energies distinguish relative binding affinities (see Table 2). The full set of data, including calculated absolute values of electrostatic solvation free energies for dielectric constants of 1, 2, 4, 8, is given as Supplemental Material. From the desolvation data, it can be observed that the free energy cost for desolvating 20E is higher than for ponA and thus disfavors binding of 20E (see Table 2). This is easily understood as 20E harbors three hydroxyl groups on its aliphatic tail, versus two for ponA. This is in complete agreement with solvation data for alcohols versus alkanes (see Table 2). The desolvation cost of the ligand is, however, compensated by the interactions made between the ligand and protein and the exact balance between these two opposite contributions has an important role in determining the resulting affinity of the ligand for the protein. More accurate computational estimates of absolute values of binding free energies entail significant computational efforts. However, in order to make a semi-quantitative estimate of the protein-ligand interaction term and avoid numerical noise from absolute value calculations, we decomposed the free energy to isolate the electrostatic contribution arising solely from the hydrogen bond between 20E and Asn504. This allowed us to focus on the contribution to binding that differs most between the two ligands given their similar chemical nature and the similarity of their interactions with the protein. As for the relative electrostatic (de)solvation free energies, the energetic contribution of the hydrogen bond between Asn504 from EcR and the hydroxyl group on C25 of 20E depends on the choice made for the dielectric constant of the ligand-protein complex. For all choices of the internal dielectric considered, the continuum electrostatic calculations resulted in a compensatory energy balance between the H-bond formed between the 25-OH group and Asn504 and the desolvation cost of the hydroxyl moiety, thus favoring ponA over 20E in terms of binding affinity (see Table 2 for data using a dielectric constant of 2; other values are given in Supplemental Material). For a more quantitative estimate of the relative affinities of ponA and 20E for their receptor (it must be recalled that the higher affinity of ponA corresponds to a difference in free energy of only 1-2 kcal/mol), detailed free energy calculations that include ligand and protein flexibility would be needed. However, the estimates afforded by the current calculations clearly stress the essential role of solvation in modulating the affinity of 20E and ponA for EcR.

Solvation Free Energy of Ecdysone and 22-deoxy-20-hydroxyecdysone. To further substantiate the importance of solvation in modulating the affinity of ligands for their cognate receptor, we estimated the desolvation free energy cost of two other ecdysteroids that bind to the ecdysone receptor. We chose ecdysone (E) and 22-deoxy-20-hydroxyecdysone (22deoxy20E) that, when compared to 20E, lack the 20-OH and the 22-OH groups, respectively. In contrast to the 25-OH group, which is placed at the end of the alkyl tail, these two hydroxyl groups are closer to the steroid core, and adjacent to each other (see scheme in Fig. 2). In contrast to the 25-OH, the 20- and 22-hydroxyl groups contribute favorably to the affinity of 20E for the receptor as judged from the lower experimental activities of E (100-fold lower) and 22deoxy20E (3-fold lower) compared to that of 20E (37). Since no x-ray crystal structure of EcR in complex with E and 22deoxy20E are available, we made the assumption that these two ligands have similar conformations and similar binding modes to the receptor, except, of course, for the specific H-bonds made between the hormone and the receptor. For example, the H-bond observed in the complex between the 20-OH of 20E and Tyr408, which could not be formed with E that lack the 20-OH group.

An important observation derived from the analysis of the 20E-bound EcR crystal structure is that the different contributions of 20-OH and 25-OH to the binding affinity cannot be simply attributed to differences in the environment/geometry of these two hydroxyl groups. In fact, similar H-bonding geometries and environments are observed for 20-OH H-bonded to Tyr408 (O-O distance 2.6 Å) and for 25-OH H-bonded to Asn504 (O-O distance 2.6 Å). These observations remained true after building in the hydrogen atoms and carrying out an energy minimization within the framework of the CHARMM force field (see Table 2 for details). Therefore, the computed estimate of the contribution of the 20-OH/Tyr408 H bond to the protein-hormone electrostatic interaction is very similar to the computed contribution for the 25-OH H-bond to Asn504 (see Table 2). For the 22-OH group, the H-bond interactions with the protein are mediated by water (see Fig. 3), and no attempt at quantitative estimates were made.
Considering then the solvation free energies for 20E, E, 22deoxy20E and ponA, the present calculations suggest a significantly lower desolvation cost for the 20- and 22-OH groups compared to that of 25-OH, as indicated in Table 2 (using $\epsilon_{\text{int}}=2$, data for all values of the dielectric constant are presented as Supplemental Data). Desolvation (which opposes binding) is larger for 20E than for ponA by about 4.5 kcal/mol (Table 2), while the desolvation cost of 20E is larger than that of E or 22deoxy20E by only about 2 kcal/mol (Table 2). This suggests that the desolvation cost of the 25-OH is larger than the desolvation cost of the 20-OH or 22-OH groups by about 2.5 kcal/mol. In order to better understand these observations, consider the molecular structure of the ligand in more detail. The 20- and 22-hydroxyl groups can be considered to form a 1,2 diol (see Table 2). Theoretical and experimental studies on 1,2 diols (48) that indicate that the desolvation cost of a 1,2 diol is not a simple sum of the desolvation cost of two hydroxyl groups. For example, the experimental solvation free energies for ethane (1.8 kcal/mol), ethanol (-5.0 kcal/mol) and 1,2 ethane diol (-9.3 kcal/mol) (49), clearly show that adjacent hydroxyl groups do not contribute additively to solvation, and that the difference between an isolated OH and an OH implicated in a 1,2 diol is about 2.5 kcal/mol. This difference can be traced to conformational preferences of simple diols, such as 1,2 ethane diol (ethylene glycol), which have been extensively studied both theoretically and experimentally (see (50) and references cited therein). These studies indicated that an isolated 1,2 ethane diol moiety is in a (gauche, gauche) conformation that allows an intra-molecular O-H …O electrostatic interaction (even if not a “classical” H-bond due to the geometrical constraints of the vicinal hydroxyl substituents). The intra-molecular stabilization afforded by this interaction is significant (51). In the presence of solvent water, there is a competition between the intra-molecular interactions and the interactions of the hydroxyl groups with water molecules. As a result, an equilibrium between (gauche, gauche) and (gauche,trans) conformations of 1,2 ethane diol is observed, but the gauche conformation is still largely prevalent (48). The present solvation free energy calculations on the ecdysteroids ligands are fully consistent with this study of 1,2 ethane diols. The intra-molecular stabilization of adjacent hydroxyl groups is observed in 20E bound to EcR, as indicated by the O20-O22 distance (2.8 Å) and the gauche conformation (O20-C20-C22-O22 dihedral angle, 50°), as well as in the crystal structures of free 20E (47) (Figure 4). In the 20E-EcR complex, the intra-molecular stabilization involves an interaction between the H atom of 20-OH and O22. This acceptor role of O22 is consistent with recent biological activity data that show that the 22-OH can be replaced by a methyl ether moiety without loss in affinity, indicating its role as an acceptor (52).

**DISCUSSION**

The 20-hydroxyecdysone (20E) is the natural ligand of the ecdysone receptor composed of the nuclear receptor heterodimer EcR/USP. This steroid hormone is the major and most active natural metabolite of the precursor ecdysone and it is present in all arthropods, from insects to crustaceans and spiders. Ecdysteroids are also present in plants and fungi, and together with animal ecdysteroids represent more than 300 different compounds (see Ecdybase at http://ecdybase.org/). The phytoecdysteroid 20E deoxygenated at the C25 position, called ponasterone A (ponA), is one of the most potent ecdysteroids found in nature, far more efficacious than the natural hormone 20E. In vitro experiments performed on cells from various species representative of different insect orders indicated that the binding affinity of ponA is much larger than that of 20E.

To better understand the structure/activity relationship of these two ligands for their receptor, the X-ray crystal structure of the ligand binding domains of the moth Heliothis virescens EcR/USP in complex with 20E was solved. The structure of EcR/USP in complex with ponA was previously solved for different insect species, the moth Heliothis (13), the silver whitefly Bemisia (14) and the beetle Tribolium (15). In their respective ligand bound ecdysone receptor crystal structures, 20E and ponA adopt almost identical chair conformations and similar positioning of their alkyl tail. The structural analysis of the ligand/receptor interactions reveal similar interactions, except for the additional hydrogen bond created between the 25-OH group of 20E and the polar residue Asn 504 in helix H11. This residue is strictly conserved in all arthropod species for which EcR has been cloned and sequenced. Given this conservation and the omnipresence of 20E among arthropods, it is most likely that the interaction of the 25-hydroxyl group of 20E and the Asn residue of H11 will be preserved across species. Despite this additional H-bond, the binding affinity of 20E for its receptor is, depending on the insect species, one to two orders of magnitude lower compared to the corresponding ponA values. In the present case, the differences cannot be simply rationalized by the structural comparison of 20E and ponA bound to EcR, and the experimental crystal structures served as starting points for a theoretical analysis based on two approaches, one employing scoring functions implemented in two popular docking programs and another approach that decomposed the protein/ligand free energy of association into specific amino-acid –ligand interactions. The results from the simple scoring functions predicted a better binding affinity for 20E compared to ponA, in contrast to the experimental observations. While having been shown to be effective in numerous docking studies, their failure here to properly predict the correct order of binding affinity suggested that other contributions, in
particular, (de)solvation may play a particularly important role in modulating the binding of these ecdysteroids. Subsequent application of the MM/PBSA free energy decomposition method, where the desolvation contribution is explicitly taken into account, indicated that the favorable contribution from the H-bond made between the Asn 504 and the 25-OH group of 20E does not completely offset the larger desolvation cost of 20E compared to ponA. Thus, the presence of the extra hydrogen bond between 20E and EcR does not lead to an increased affinity for the ligand 20E. These results are consistent with experimental studies comparing 20E versus ponA binding. The crucial contribution of solvation and desolvation to binding affinities has proved to be an increasingly relevant point of consideration. Numerous examples, where these contributions play an important role, have been discussed in the literature. In particular, in the problematic of carbohydrate-protein recognition, the role of water was shown to be essential in modulating the binding affinity of carbohydrates to the protein, even though carbohydrate hydroxyl groups that bind to the protein do not significantly contribute to the binding affinity (53). Some similarities exist between the carbohydrate binding to proteins and the binding of 20E to its cognate nuclear receptor. In fact, the various hydroxyl groups of 20E, and in particular those at positions C20, C22 and C25 (see Table 2), play different roles in modulating the affinity of the ligand for the receptor. The present free energy analysis indicates that the differential role of 20-, 22-, and 25-OH to the complex stability resides mostly in their different contributions to solvation. In fact, the calculated relative solvation free energies for 20E, E, 22deoxy20E and ponA (Table 2) indicate that the desolvation costs of the 20- and 22-hydroxyl groups are much lower than that of the 25-OH group. Consistent with this are results of CoMFA data that suggest that the 20- and 22-OH groups have stabilizing effects on ligand-receptor complex formation, while the 25-OH group has a destabilizing contribution (37). As a consequence, even though E, 22deoxy20E and ponA each have two hydroxyl groups in their aliphatic tail (see Table 2), ponA is poised to have an increased affinity for the receptor with respect to ligands such as E or 22deoxy20E, because the adjacent hydroxyl in ponA have a lower desolvation cost than the separate hydroxyls of E or 22deoxy20E. The adjacent 20- and 22-OH moieties form a 1,2 diol that is stabilized by intra-molecular interactions between the adjacent hydroxyl groups. Interestingly, the 2-β and 3-β OH groups also form a 1,2 diol and thus will have a lower desolvation cost with respect to non adjacent hydroxyl. In this case, there is an even more striking resemblance to the adjacent hydroxyl substitutions seen in carbohydrate rings. Diol-protein interactions thus appear to be a simple structural motif well recognized in protein-carbohydrate complexes, but that is also used in other protein-ligand complexes, such as nuclear receptor-ligand complexes. These types of interactions exploit simple physico-chemical principles of intra-molecular versus inter-molecular stabilization to confer larger stabilizing effects to particular H-bond interactions.

Therefore, despite the simplified description of solvation afforded by the continuum electrostatics model, a coherent picture emerges from the ensemble of data and points to quantitative differences in the solvation free energy of the adjacent hydroxyl groups (giving rise to 1,2 diol for 20- and 22-OH) and the non-adjacent hydroxyl group (25-OH in 20E) at the origin of the large difference in affinity observed for 20E and ponA. The crucial role played by solvation is further emphasized by the comparative analysis with free 20E. In the dehydrated free 20E structure, the 20-, 22- and 25-OH groups are linked together through H-bonds. In the hydrated states, the 20- and 22-OH groups maintain their conformations and form a diol bond, while a change in the position of 25-OH is observed that results in a conformation of the alkyl tail similar to that observed for 20E bound inside the EcR LBP (see Fig. 4). Thus, these crystallographic data further support the importance of water solvation at the 25-OH position, while the conformations of the 20- and 22-hydroxyl groups are less sensitive to solvation. In fact, from solvation data alone, the added cost for desolvating 25-OH implies that even if it forms a hydrogen bond to the protein that is as stable as that formed between 20-OH and the protein, its contribution to the complex stability should be two to three orders of magnitude less than that of the 20-OH group.

REFERENCES

1. Gilbert, L. I., Rybczynski, R., and Warren, J. T. (2002) Annu. Rev. Entomol. 47, 883-916
2. Gilbert, L. I. (2004) Molecular and Cellular Endocrinology 215, 1-10
3. Koelle, M. R., Talbot, W. S., Segraves, W. A., Bender, M., Cherbas, P., and Hogness, D. S. (1991) Cell 67, 59-77
4. Thomas, H. E., Stunnenberg, H. G., and Stewart, A. F. (1993) Nature 362, 471-475
5. Yao, T.-P., Segraves, W. A., Oro, A. E., McKeown, M., and Evans, R. M. (1992) Cell 71, 63-72
6. Yao, T.-P., Forman, B. M., Jiang, Z., Cherbas, L., Chen, J.-D., McKeown, M., Cherbas, P., and Evans, R. M. (1993) Nature 366, 476-479
46. Hu, X., Cherbas, L., and Cherbas, P. (2003) Mol.Endocrinol. 17, 716-731
47. Fabian, L., Argay, G., Kalman, A., and Bathori, M. (2002) Acta Cryst.B 58, 710-720
48. Cramer, C. J. and Truhlar, D. G. (1994) J.Am.Chem.Soc. 116, 3892-3900
49. Rizzo, R. C., Aynechi, T., Case, D. A., and Kuntz, I. D. (2006) J.Chem.Theory Comput 2, 128-139
50. Crittenden, D. L. (2005) J.Phys.Chem.A 109, 2971-2975
51. Trindle, C., Crum, P., and Douglass, K. (2003) J.Phys.Chem.A 107, 6236-6242
52. Lapenna, S., Hormann, R. E., and Dinan, L. (2007) J.Insect Science 7:13. in Bellés X, Billas I, Cherbas P, Delbecque P, Douglass, K. Abstracts of the 16th International Ecdysone Workshop. 57pp, 32-33
53. Lemieux, R. U. (1996) Acc.Chem.Res. 29, 373-380

FOOTNOTES

We thank the Structural Biology and Genomics Platform at IGBMC for their help at various stages of this work; V. Chavant for technical assistance; A. Mitschler for help during data collection; the staff of ESRF ID14 (Grenoble, France) and of the SLS (Villigen, Switzerland) for assistance during synchrotron data collection; IDRIS and CINES for generous allowance of computer time; CNRS, INSERM and the University Louis Pasteur for support and R.A. Atkinson for careful reading of the manuscript. The work was supported in part by Bayer CropScience, the Association pour la Recherche sur le Cancer and the European Commission SPINE project (Contract No QLG2-CT-2002-00988) under the Integrated Programme “Quality of Life and Management of Living Ressources”. EM acknowledges financial support from the Ligue Contre le Cancer.

The abbreviations used are: EcR: ecdysone receptor; USP: Ultraspiracle protein; 20E: 20-hydroxyecdysone; E: ecdysone; 22deoxy20E: 22-deoxy-20-hydroxyecdysone; ponA: ponasterone A; DBH: dibenzoylhydrazine; LBD: ligand binding domain; LBP: ligand binding pocket

The atomic coordinates and structure factors (code 2R40) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

FIGURE LEGENDS

Figure 1. A, Overall structure of the ligand-binding domains of Heliothis EcR/USP bound to its endogenous ligand 20-hydroxyecdysone (20E) with HvEcR in yellow and HvUSP in green. Helices H12 are shown in red. The ligands 20E in the ligand-binding pocket of EcR and the phospholipid inside USP are represented by sticks coloured by atom types (grey for carbon and red for oxygen). The sigma-A-weighted 2Fo-Fc electron density map is shown at the level of 20E where it is contoured at 1.0σ and overlaid on the final refined model as a cyan mesh. B, Enlarged view of the region of EcR comprising helices H1, H3, H9 and H12 where additive molecules are seen at the solvent-exposed surface of the receptor. A glycerol molecule (gly) is located in the coactivator binding cleft and forms a water-mediated H-bond to Q524 (H12), while a citrate molecule and a sulfate ion are located in the cleft created delimited by H1, H3 and the loop connecting H3 to H4. The sigma-A-weighted 2Fo-Fc electron density map overlaid on the additive molecules is shown at a contour level of 1.0σ. The additive molecules are shown in stick representation coloured by atom type (grey for carbon, red for oxygen and yellow for sulphur). The EcR residues are shown by a stick representation with yellow for carbon, red for oxygen and blue for nitrogen.

Figure 2. A, Schematic representation of 20-hydroxyecdysone (20E). B, 20E in the ligand-binding pocket of its cognate Heliothis EcR receptor. The EcR amino acid residues and 20E are shown in stick representation coloured by atom type (light blue and wheat for carbon, respectively, red for oxygen, blue for nitrogen and yellow for sulphur). Water molecules are represented by red balls. Hydrogen bonds between ligand and residues are indicated by blue dotted lines and the green dotted line indicates a close van der Waals contact between the 25-OH group of 20E and the Cζ3 of W526 (distance 3.0 Å).

Figure 3. Stereoview of the water-mediated interaction network in the region of the 20E alkyl side chain. The H-bond interaction network involves the 20-, 22- and 25-OH groups of the 20E alkyl side chain and three structural water molecules. The Cα-trace of the EcR protein backbone is shown in a yellow ribbon representation, except for H12 which is shown by a red Cα-trace. The EcR amino acid residues and 20E are shown in stick representation coloured by atom type (light blue and wheat for carbon, respectively, red for oxygen and blue for nitrogen). Water molecules are represented by red balls. Hydrogen bonds between ligand
and residues are indicated by blue dotted lines and the green dotted line indicates a close van der Waals contact between the 25-OH group of 20E and the C3 of W526 (distance 3.0 Å).

**Figure 4.** X-ray crystal structures of 20E bound to its cognate ecdysone receptor and the free 20E molecule. **A,** Structural conformation of 20E bound to *Heliothis* EcR showing in more detail the conformation of the 20E alkyl side chain and the H-bond interaction network (yellow dotted lines) of the 20-, 22- and 25-OH groups with protein neighbouring residues and the structural water molecules. The 20E ligand and the amino acid residues are represented by sticks coloured in wheat and light blue, respectively for carbon and in red for oxygen. **B-C,** Crystal structures of the 20E molecule crystallized in a solvent-free conformation (**B**) and in a trihydrated conformation (**C**) (47). 20E is shown by sticks coloured in cyan (**B**) and grey (**C**) for carbon and in red for oxygen. Water molecules are shown by magenta balls and fragments of symmetrical 20E molecules in (**C**) by grey and red sticks for carbon and oxygen, respectively.

**Figure 5.** Electrostatic (ELEC: white) and van der Waals (VDW: black) contributions to the interaction energy between the ligand 20E and the LBD of HvEcR. The interactions are decomposed into individual interactions from the LBD amino acids, see text for details. Electrostatic interactions are estimated with continuum electrostatics using a dielectric constant of 2 for the protein and 80 for the solvent. Only amino acids with significant contributions discussed in the text are represented. See text for details and Supplemental Data for a full list of contributing amino acids. Units: kcal/mol.
### Table 1. Crystallographic data and refinement statistics

| Data Processing          |                  |
|--------------------------|------------------|
| Synchrotron              | SLS (Villigen, CH) |
| Resolution (Å)           | 50.0-2.4         |
| Crystal space group      | P3,21            |
| Cell parameters (Å)      |                  |
|                          | a = b = 58.001   |
|                          | c = 303.592      |
|                          | x = y = 90°      |
|                          | z = 120°         |
| Total reflections        | 210445           |
| Unique reflections       | 24160            |
| Mean redundancy          | 8.7              |
| Rsym (%) (last shell)*   | 5.0 (39.4)       |
| Completeness (%)         | 99.2 (99.7)      |
| Mean I/σ (%) (last shell)| 35.29 (4.81)     |

| Refinement               |                  |
|--------------------------|------------------|
| Number of protein atoms  | 467              |
| Number of solvent sites  | 87               |
| Number of EcR ligand atoms | 34            |
| Number of USP ligand atoms | 49             |
| R.m.s.d. bond length (Å)¶ | 0.01            |
| R.m.s.d bond angles (°)¶ | 1.289           |
| R.m.s.d improper dihedrals (°)¶ | 0.567         |
| Rcryst (Rfree) (%)§      | 22.39 (28.18)    |

Values in parentheses correspond to the highest resolution shell

*Rsym (I) = \( \Sigma_{hkl} \Sigma_i | I_{hkl,i} - < I_{hkl} > | / \Sigma_{hkl} \Sigma_i | I_{hkl,i} | \) with < I_{hkl} > the mean intensity of the multiple I_{hkl,i} observations for symmetry-related reflections.

¶Root-mean-squared deviation (R.m.s.d.) are given from ideal values.

§Rcryst = \( \Sigma_{hkl} | F_{obs} - F_{calc} | / \Sigma_{hkl} | F_{obs} | \) where F_{obs} and F_{calc} are the observed and calculated structure amplitudes, respectively. R_{free} is the same as R_{cryst} but calculated on the 10% of data excluded from refinement.
Table 2.
Relative binding affinities for the ecdysone receptor between 20-hydroxyecdysone (20E) and the ecdysteroids ponasterone A (ponA), ecdysone (E) and 22-deoxy-20-hydroxyecdysone (22deoxy20E). For sake of clarity, the aliphatic tails of the four ligands are represented. Negative values indicate that the free energy (or free energy contribution) favors the ligand considered over 20E, while positive values indicate that 20E is favored over the ligand considered. Units: kcal/mol.

$\Delta \Delta G^{\text{elec}}_{\text{desolv}}$ _EXP_ : Experimental estimates for the contribution of desolvation to the relative binding affinities. Obtained by comparison of the experimental solvation free energies (Rizzo, 2006) for ethane (1.8 kcal/mol) and ethanol (-5.0 kcal/mol) (comparison 20E/ponA) and from those of ethanol and 1,2 ethane diol (-9.3 kcal/mol) (comparison E or 22deoxyE and 20E).

$\Delta \Delta G^{\text{elec}}_{\text{desolv}}$ _CALC_ : Calculated electrostatic contribution to the relative desolvation are estimated for the full ecdysteroid ligands as described in Mat & Met using an internal dielectric constant of 2.

$\Delta \Delta G^{\text{elec}}_{H-bond}$ _CALC_ : Calculated relative affinity contribution due to formation of H-bond between the ligand and protein. PonA: contribution of the H-bond made by 25-OH is lost, E contribution of the 20-OH H bond is lost. Calculations based on the experimental structure for the HvEcR-20E complex. In the absence of experimental structure for the E-bound HvEcR complex, the structure was assumed to be similar to that of the 20E-bound HvEcR. The H-bond made by 22-OH is water-mediated, hence no calculation of its contribution was attempted with continuum electrostatics methods. After energy minimization, the geometrical parameters for the 25-OH – N504 H-bond are: distance H…O: 1.92 Å); angle O-H-O: 171°; angle C-O-H 134° while for the (20-OH - Y408 H-bond they are 1.88 Å, 163° and 127°, respectively). See text for details.

$\Delta \Delta G^{\text{elec}}_{\text{tot}}$ _CALC_ : Total calculated electrostatic contribution to the relative affinity, i.e. sum of the desolv and H-bond terms above.

$\Delta \Delta G^{\text{elec}}_{\text{tot}}$ _EXP_ : Experimental relative free energies as estimated from the experimental values of -log(ED$_{50}$) for DmEcR (37).

|         | 20E       | PonA     | E         | 22deoxy-20E |
|---------|-----------|----------|-----------|-------------|
| $\Delta \Delta G^{\text{elec}}_{\text{desolv}}$ _EXP_ | 0         | -6.8$^{(a)}$ | -4.3$^{(b)}$ | -4.3$^{(b)}$ |
| $\Delta \Delta G^{\text{elec}}_{\text{desolv}}$ _CALC_ | 0         | -4.3     | -1.9      | -1.8        |
| $\Delta \Delta G^{\text{elec}}_{H-bond}$ _CALC_ | 0         | 4.4      | (+4.3)    | -----       |
| $\Delta \Delta G^{\text{elec}}_{\text{tot}}$ _CALC_ | 0         | 0.1      | (+2.4)    | -----       |
| $\Delta \Delta G^{\text{tot}}_{\text{EXP}}$ | 0         | -1.4     | 2.2       | 0.3         |

(a) Estimated from of the experimental solvation free energies for ethane and ethanol (b) for ethanol and 1,2 ethane-diol. (see legend)
Figure 1
Figure 2
Figure 4
Figure 5