Structure of 3α-Hydroxysteroid/Dihydrodiol Dehydrogenase Complexed with NADP+ †

Melanie J. Bennett,‡ Brian P. Schlegel,§ Joseph M. Jez,‡ Trevor M. Penning,§ and Mitchell Lewis*,‡

Department of Biochemistry and Biophysics, The Johnson Research Foundation, and Department of Pharmacology, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104-6059

Received February 27, 1996; Revised Manuscript Received May 28, 1996⊗

ABSTRACT: Rat liver 3α-hydroxysteroid/dihydrodiol dehydrogenase (3α-HSD) inactivates circulating steroid hormones and is involved in polycyclic aromatic hydrocarbon (PAH) carcinogenesis. It is the only HSD of known structure in the aldo−keto reductase (AKR) superfamily and may provide a paradigm for other mammalian HSDs in this family. The structure of the 3α-HSD·NADP+ binary complex has been determined at 2.7 Å resolution and refined to a crystallographic *R*-factor of 23.4% with good geometry. The model is similar to other binary complexes in the AKR superfamily in that NADP+ binds at the C-terminal end of an α/β barrel. However, it is unique in that NADP+ is bound in two alternate conformations, probably because of the lack of a salt-linked "safety belt" over the pyrophosphate bridge. The structure supports a previously proposed catalytic mechanism for carbonyl reduction in which Tyr 55 is the general acid, and its effective pK_a is lowered by the adjacent Lys 84. We present evidence that the structurally distinct short-chain dehydrogenase/reductase (SDR) superfamily may have convergently evolved a similar catalytic mechanism. Insight into substrate binding is offered by a crystal packing contact in which a neighboring molecule inserts a tryptophan residue (Trp 227) into an apolar cleft in 3α-HSD. This cleft is proximal to the bound NADP+ cofactor and contains a surface of apolar residues (Leu 54, Trp 86, Leu 122, Phe 128, Phe 129, Leu 137, Phe 139), making it a likely candidate for the substrate-binding site. Thus, in forming this crystal contact, Trp 227 may mimic a portion of a bound steroid. In addition, we propose that a water molecule in the active site indicates the position of the hydroxyl oxygen in a 3α-hydroxysteroid substrate. Knowledge of the position of this water molecule, combined with the stereochemistry of hydride transfer, suggests that the α face of a bound steroid will be oriented toward the side of the apolar cleft containing Trp 86.

Hydroxysteroid dehydrogenases (HSDs)[1] play pivotal roles in the biosynthesis of all steroid hormones including glucocorticoids, mineralocorticoids, androgens, progestins, and estrogens. In steroid hormone target tissues they reversibly inactivate steroid hormones by interconverting a single carbonyl/hydroxyl group. In this way, HSDs act as molecular switches by regulating the binding of steroid hormones to their receptors, ultimately modulating the transcription of target genes in the nucleus. HSDs belong to at least two distinct protein families: the short-chain dehydrogenase/ reductase (SDR) and aldo−keto reductase (AKR) super-families (Jornvall et al., 1995; Bruce et al., 1994). HSDs belonging to these two groups interconvert similar substrates but have very different three-dimensional structures. Comparison of SDR and AKR enzymes may provide insight into their structural/functional relationships.

Rat liver 3α-hydroxysteroid/dihydrodiol dehydrogenase (3α-HSD) (EC 1.1.1.213)[2] is a monomeric oxidoreductase with dual NAD(P)(H) cofactor specificity and a molecular weight of 37 029 (Pawlowski et al., 1991). It was the first HSD to be assigned to the AKR superfamily and is known to reduce biologically active 3-ketosteroids of the androstane (C19), pregnane (C21), and cholane (C24) series (Mukharji, 1982). These reactions lead to the inactivation of circulating androgens, progestins, and glucocorticoids, as well as the biosynthesis of bile acids. 3α-HSD also functions as a dihydrodiol dehydrogenase (DD). In this role, it oxidizes polycyclic aromatic hydrocarbon (PAH) *trans*-dihydrodiols (proximate carcinogens) to reactive *o*-quinones (Smithgall et al., 1988) with the concomitant production of reactive oxygen species and semiquinone anion radicals which may contribute to PAH carcinogenesis (Penning, 1993; Penning et al., 1996).

3α-HSD is the only HSD of known structure in the AKR superfamily. Human and porcine aldose reductase were the first AKR structures determined and revealed the superfamily fold to be an (α/β)$_8$ barrel (Wilson et al., 1992; Rondeau et al., 1992). Rat liver 3α-HSD shares 58% sequence identity with human aldose reductase, which permitted the crystal-

* Address correspondence to this author at the Department of Biochemistry and Biophysics, University of Pennsylvania School of Medicine, 37th and Hamilton Walk, Philadelphia, PA 19104-6059. Telephone: (215) 898-0949. Fax: (215) 898-4217. E-mail: lewis@xtal. med.upenn.edu.
‡ Department of Biochemistry and Biophysics.
§ Department of Pharmacology.
⊗ Abstract published in *Advance ACS Abstracts,* August 1, 1996.
[1] Abbreviations: AKR, aldo−keto reductase; DD, dihydrodiol dehydrogenase; HSD, hydroxysteroid dehydrogenase; NAD+, nicotinamide adenine dinucleotide (oxidized form); NADH, nicotinamide adenine dinucleotide (reduced form); NADP+, nicotinamide adenine dinucleotide phosphate (oxidized form); NADPH, nicotinamide adenine dinucleotide phosphate (reduced form); PAH, polycyclic aromatic hydrocarbon; SDR, short-chain dehydrogenase/reductase.

[2] Formerly EC 1.1.1.50; however, new conventions to distinguish A and B class dehydrogenases stipulate a change to EC 1.1.1.213.

lographic structure determination of apo 3α-HSD by molecular replacement at 3.0 Å resolution (Hoog et al., 1994). Now, we have determined the structure of the 3α-HSD·NADP⁺ binary complex at 2.7 Å resolution.

cDNAs and deduced amino acid sequences for other HSDs reveal that many belong to the AKR superfamily. These include human prostate 3α-HSD which is involved in androgen metabolism (Lin et al., 1995; Dufort et al., 1995), human DD2 and DD4 which also function as 3α-HSDs (Deyashiki et al., 1994; Stolz et al., 1993; Khanna et al., 1995), murine liver 17β-HSD (Deyashiki et al., 1995), and rat and rabbit ovarian 20α-HSDs (Miura et al., 1994; Lacy et al., 1993). On the basis of the high sequence identity that exists between HSDs in the superfamily, the 3α-HSD·NADP⁺ binary complex may provide insight into cofactor and substrate binding and catalysis in these other mammalian HSDs.

The three-dimensional structure of the binary complex reveals the mode of cofactor binding and provides a structural explanation for some observations on the stereochemistry of the reaction. Previous mechanistic studies established that 3α-HSD is a class A dehydrogenase that transfers the *pro-R* hydrogen from C4 of the nicotinamide ring to the acceptor carbonyl of a 3-ketosteroid substrate. The structure explains the 4-*pro-R* hydrogen specificity in terms of the amino acid side chains interacting with the nicotinamide ring. 3α-HSD also displays a sequential ordered kinetic mechanism in which pyridine nucleotide binds first and leaves last and in which binding and release of cofactor is rate limiting for oxidation as well as reduction, implying that a conformational change may occur on binding NAD(P)(H) (Askonas et al., 1991). Comparison of the apo and binary complex 3α-HSD structures reveals some minor changes in side chain positions upon cofactor binding. However, a loop (loop B), which was expected to undergo a large conformational change based on similarity to other AKR structures, is involved in a crystal contact, and we cannot determine whether there is a change in this loop upon cofactor binding.

In addition to revealing features of cofactor binding, the binary complex model provides insight into substrate recognition and catalysis. First, the C4 position of the nicotinamide ring lies at the base of an apolar cleft which is thus implicated in substrate binding. A crystal packing contact in this cleft may mimic some features of steroid binding. Second, a water molecule forming hydrogen bonds with Tyr 55 and His 117 may indicate the position of the hydroxyl or carbonyl oxygen in a substrate, which allows us to propose a model for the orientation of a bound steroid based on stereochemical constraints. Third, analysis of the three-dimensional structure combined with mutagenesis data supports a previously proposed catalytic mechanism which involves Tyr 55 and Lys 84 (Hoog et al., 1994). Finally, superposition of the active site tyrosine/lysine pair in 3α-HSD and 3α,20β-HSD (a SDR superfamily member) suggests that these functionally similar, yet structurally distinct enzymes may have convergently evolved a common reaction mechanism.

EXPERIMENTAL PROCEDURES

Purification and Crystal Growth. Homogeneous 3α-HSD from male Sprague-Dawley rat livers was prepared using the published procedure (Penning et al., 1984). Aliquots of enzyme were stored at −70 °C in a glycerol-containing buffer. Before use, enzyme was thawed and dialyzed against three changes of buffer (10 mM potassium phosphate, pH 7.0, 1 mM EDTA, 1 mM 2-mercaptoethanol). Dialyzed enzyme was concentrated to 10 mg/mL, and the binary complex was prepared by adding a 10-fold molar excess of NADPH. It was later determined by absorption spectroscopy at 260 and 340 nm that, under the crystallization conditions, which involve acidic pH, NADPH is oxidized to NADP⁺.

Binary complex crystals were grown by hanging drop vapor diffusion with seeding from 27% (w/v) poly(ethylene glycol) (PEG) 4000, 0.19 M ammonium acetate, 0.10 M sodium acetate, pH 4.6 (final pH = 5.7), and 1.3% (v/v) dimethyl sulfoxide. Hanging drops containing 10 mg/mL binary complex produced only multiple crystals which were unsuited to X-ray analysis. Hanging drops containing 5 mg/mL binary complex produced small single crystals with good morphology after streak seeding. These small crystals were washed in 50% mother liquor and used as macroseeds in fresh crystallization drops. The macroseeded crystals grow to a size of 0.50 mm × 0.25 mm × 0.05 mm and belong to space group $P2_1$ with unit cell dimensions $a = 46.9$ Å, $b = 139.1$ Å, $c = 53.3$ Å, and $\beta = 113.5°$. Assuming two molecules per asymmetric unit, the Matthews number (Matthews, 1968) is 2.2 Å³/Da, consistent with the finding of two independent cross-rotation function solutions (discussed below).

X-ray Data Collection. All data were collected at room temperature. Initially, data to 3.2 Å were collected from one crystal (0.20 mm × 0.20 mm × 0.05 mm) using a Siemens area detector. Data were collected for 45 s using 0.17° ϕ scans at a crystal to detector distance of 140 mm with $2\theta = 10°$ and $\lambda = 1.54$ Å. The data were integrated and scaled using the programs XDS and XSCALE (Kabsch, 1988). A total of 28 823 reflections were merged to give 9728 out of 10 274 possible unique reflections at 3.2 Å with $R_{merge} = 0.079$. Higher resolution data to 2.7 Å were later collected from one binary complex crystal (0.50 mm × 0.25 mm × 0.05 mm) using an RAXIS-II image plate. Oscillation frames were collected from two positions on the crystal using a rotation of 1.5° for 25 min at a crystal to detector distance of 110 mm and a wavelength of 1.54 Å. The data were integrated and scaled using MOSFLM (Leslie, 1992) and the CCP4 suite (1994). A total of 53 539 reflections were merged to give 16 123 out of 17 221 possible unique reflections at 2.7 Å resolution with $R_{merge} = 0.058$.

Molecular Replacement. The binary complex structure was determined by molecular replacement using the programs in X-PLOR (Brunger, 1992a). The 3.0 Å resolution apo 3α-HSD model (Hoog et al., 1994) consisting of residues 1−308 provided the search model in a cross-rotation function using data in the resolution range 15−4 Å. The top peaks in the rotation function were 7.8σ and 7.4σ above the mean and were related by a noncrystallographic 2-fold rotation axis in the x−z plane, consistent with the 222 symmetry observed in the self-rotation function ($\Psi = 90.0$, $\Phi = 90.00$, $\kappa = 180.0$; $\Psi = 90.0$, $\Phi = 0.00$, $\kappa = 180.0$).

A translation search was carried out in two dimensions (x and z). The first molecule was positioned using the top rotation function solution, and a translation function was calculated using data in the resolution range 15−4 Å. A single peak was observed of 5.8σ above the mean. The second molecule was oriented using the noncrystallographic

symmetry observed in the self-rotation function, and a translation function was calculated to position it in the $x-z$ plane. The top solution for this oriented molecule was 8.7σ above the mean. The relative y value was established by translating the second molecule relative to the first, producing a solution with a peak height 6.8σ above the mean.

Modeling and Refinement of the Model to X-ray Data. Refinement was performed using the programs in X-PLOR (Brunger, 1992a) with strict noncrystallographic symmetry (NCS) constraints imposed for all but the final round. Atomic coordinates obtained from molecular replacement were refined against the 3.2 Å data set for 20 cycles with the entire molecule modeled as a rigid body, resulting in a crystallographic R-factor (R_{cryst}) of 42.4%. Prior to atomic refinement using the 2.7 Å data set, several loops (residues 1–6, 131–137, and 230–234) were deleted from the model because of close crystal contacts. A typical round of refinement consisted of conjugate gradient minimization and simulated annealing refinement against 12 308 observed structure factor amplitudes between 5 and 2.7 Å. A subset of 1386 structure factor amplitudes was used to monitor the free R-factor (R_{free}) (Brunger, 1992b). After 50 initial cycles of conjugate gradient minimization, R_{free} was 40.0% and R_{cryst} was 32.8%. Conventional difference maps were calculated, and residues 1–6 were rebuilt, NADP$^+$ was modeled into electron density in the active site, and protein side chain and main chain atoms were fit, resulting in $R_{free} = 34.4\%$ and $R_{cryst} = 25.4\%$. Several rounds of rebuilding and refinement, including the modeling of an alternate conformation for NADP$^+$ (which was indicated by difference density in both molecules in the asymmetric unit), decreased R_{free} and R_{cryst} to 29.4% and 23.4%, respectively. However, the electron density was still poor for portions of the two extended loops A and B (residues 131–136 and 221–233). To further improve the quality of the phases, the electron density was 2-fold averaged and solvent flattened, which permitted modeling of the loops and confirmed the NADP$^+$ model.

Molecular Averaging and Solvent Flattening. Two-fold molecular averaging and solvent flattening was done with programs in the PHASES package (Furey & Swaminathan, 1990). Molecular masks for averaging and solvent flattening included the core of the molecule (residues 1–26, 30–130, 137–220, and 234–308), which is identical in the two molecules in the asymmetric unit related by 2-fold NCS. Extended loop regions which had the potential to deviate from the NCS were included in molecular masks for solvent flattening but not averaging (residues 131–136 and 221–233). Initial phases were calculated from the core model, not including the coordinates for NADP$^+$ and loop residues 131–136 and 221–233. Five iterations of solvent flattening and seven iterations of alternating averaging and solvent flattening using F_o coefficients converged with an R-factor between F_o and F_{av} (the result of back transformation of the averaged map) of 13.2%.

Modeling and Refinement of Loops. The electron density produced by averaging and solvent flattening was very good for the cofactor, and somewhat improved for the loops, permitting residues 225–233 and 131–136 to be modeled. Even so, residues 222–224 were still disordered. In addition, a distinct peak of electron density appeared in the active site and was modeled as a water molecule. An additional round of refinement was carried out using loose NCS restraints for the loops that differ in the two molecules (residues 132–

136 and residues 225–233), while maintaining tight restraints on all other residues in the model. The model was best fit to the averaged density by refining against the real and imaginary components, resulting in a vector R-factor of 43.2%. Because of the tight NCS restraints, the two molecules in the asymmetric unit are virtually identical, with a root-mean-square difference of 0.005 Å for all non-hydrogen atoms excluding loop residues 132–136 and 225–233. The final model contains 305 (residues 1–26, 30–221, and 225–311) out of 322 residues in the protein. Because of side chain disorder, two segments were modeled as polyalanine (molecule 1, 228–229; molecule 2, 228–231).

RESULTS AND DISCUSSION

Binary Complex Structure and Similarity to Apo 3α-HSD. The structure of the 3α-HSD·NADP$^+$ binary complex is shown schematically in Figure 1. As expected, the protein fold of the binary complex is similar to that of apo 3α-HSD (Hoog et al., 1994). The protein chain forms an α/β barrel, with a cylindrical core of eight parallel β-strands surrounded by eight α-helices running antiparallel to the strands. The connections between α-helices and β-strands in the barrel motif are short loops of less than 10 residues with four exceptions: two helices at the C-terminal end of the barrel (H1, residues 239–248; H2, residues 290–298) and two loops protruding from the C-terminal end of the barrel (loop A, residues 117–143; loop B, residues 217–238) (Figure 1b). A β-hairpin seals the N-terminal end of the barrel (residues 7–17), while the C-terminus is a 23-residue coil, of which the last 11 residues are disordered. Although there are nine cysteine residues, there are no disulfide bonds. NADP$^+$ binds at the C-terminal end of the barrel, perpendicular to the barrel axis, with the adenine ring at the periphery and the nicotinamide ring toward the core of the barrel.

The protein backbones of apo 3α-HSD and the binary complex are very similar, with an overall rms difference between Cα atoms of 0.7 Å, excluding 24 residues with rms differences >2 Å (residues 1–5, 25–33, 133–136, 225–232, 250–251, and 307–308). These residues are distant from the cofactor binding site and near crystal packing contacts, so their movements are probably unrelated to NADP$^+$ binding. The largest change caused by crystal packing is at the N-terminus, where residues 1–6 form a coil preceding the first strand in the β-hairpin, and move as much as 23 Å relative to the apoenzyme structure. Although all of the large main chain movements in the binary complex appear to be caused by crystal packing, several side chains undergo smaller movements that can be attributed to NADP$^+$ binding (discussed below).

Quality of the Model. The quality of the refined binary complex model can be assessed from the statistics given in Table 1. All data between 5 and 2.7 Å were used in the refinement, yielding a crystallographic R-factor of 23.4% for 12 308 reflections and a free R-factor of 29.4% for 1386 reflections. There are two molecules in the asymmetric unit related by 2-fold noncrystallographic symmetry. The refined model includes a single NADP$^+$ molecule per protein molecule, which exists in two alternate conformations with relative occupancies of 65% and 30%, and a single water molecule per protein molecule, which was modeled into
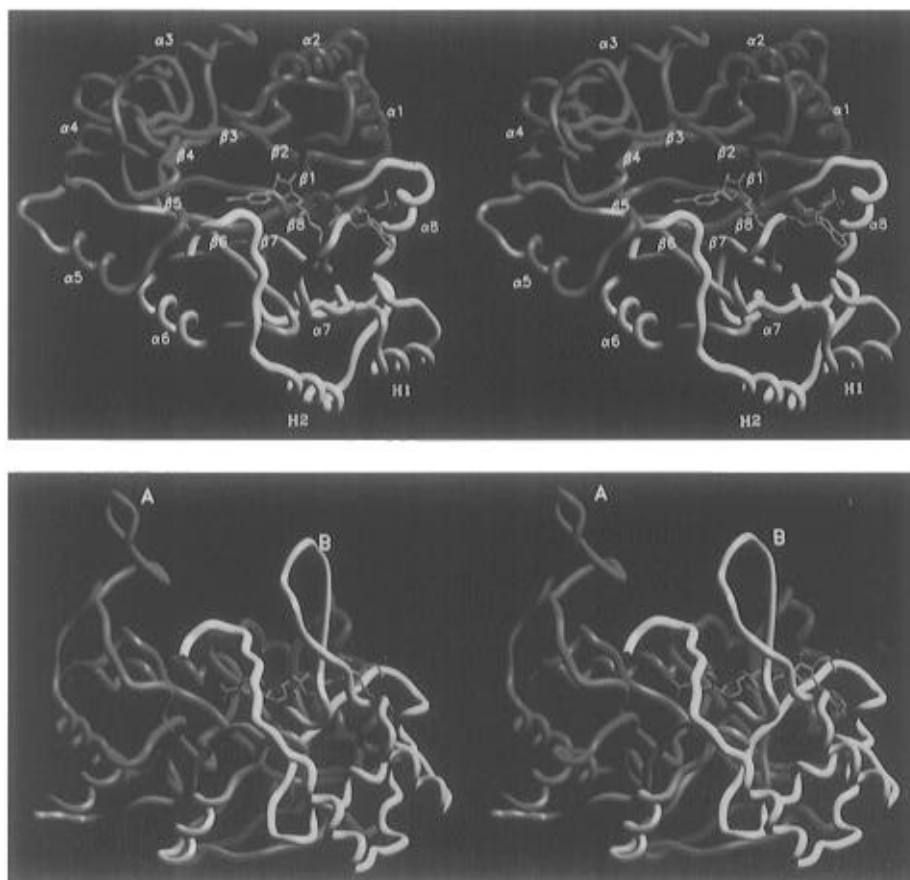
FIGURE 1: 3α-HSD·NADP⁺ binary complex structure. (a, top) Stereo figures of the 2.7 Å resolution model of the binary complex with a schematic representation of the protein backbone (top view). The color varies from dark blue (N-terminus) to red to white (C-terminus). All non-hydrogen atoms of the high occupancy conformation of NADP⁺ are shown in green. Secondary structural elements are labeled. The residue numbers in each segment are as follows: β-hairpin 1, 7- 9; β-hairpin 2, 15−17; β1, 20−22; α1, 32−43; β2, 48−50; α2, 58−70; β3, 80−85; α3, 95−106; β4, 111−116; α4, 144−156; β5, 160−166; α5, 170−177; β6, 188−192; α6, 200−209; β7, 212−216; H1, 239−248; α7, 252−262; β8, 265−269; α8, 274−280; H2, 290−298. (b, bottom) Stereo figures of the binary complex model (side view). The view is rotated 90° about the horizontal axis relative to panel a. Loops A (residues 117−143) and B (residues 217−238) are labeled. Although chain breaks are not shown, portions of two loops (residues 27−29 and 222−224) are disordered, as are the C-terminal residues 312−322. All figures except 2, 5, 8, and 9 were prepared using SETOR (Evans, 1993).

Table 1: Quality of the 3α-HSD·NADP⁺ Binary Complex Structure

free *R*-factor (%)[a]	29.4
crystallographic *R*-factor (%)[a]	23.4
resolution range (Å)	5−2.7
RMS deviations from target geometry	
bond lengths (Å)	0.012
bond angles (deg)	1.7
dihedral angles (deg)	25.2
improper angles (deg)	1.5
no. of non-hydrogen atoms[b]	
protein	4906
NADP⁺	192[c]
water	2
total	5100
single overall *B*-factor (Å²)	29

[a] $R\text{-factor} = \sum_h ||F_o(h)| - |F_c(h)||/\sum_h |F_o(h)|$, where F_o and F_c are the observed and calculated structure factor amplitudes for the reflection with Miller indices $h = (h,k,l)$. The free *R*-factor is calculated for a "test" set of reflections which were not included in atomic refinement (Brunger, 1992b). [b] Two molecules in the asymmetric unit. [c] Two conformations of NADP⁺ were modeled in each of the two active sites, one with 65% occupancy and the other with 30% (see text).

difference density in the active site. Of the 322 amino acids in the protein, 305 were modeled into the electron density. Portions of two loops (residues 27−29 and 222−224) and the C-terminus (residues 312−322) were classified as disordered, and these residues were not included in the

model. A single overall temperature (*B*) factor of 29 Å² was fit to the data.

The binary complex model compares favorably with other structures at this resolution; 88% of the backbone dihedral angles of the refined model are in the most favored regions of the Ramachandran plot as defined by the stereochemistry checking program PROCHECK (Laskowski et al., 1993). The model also fits the electron density maps well, as assessed by manual inspection and real space correlation coefficient plots (Jones et al., 1991). Although there are several residues for which the real space correlation coefficient indicates relative disorder, virtually all are in loops or chain termini which are expected to be more mobile.

In addition, the model has generally high 3D−1D profile scores (Bowie et al., 1991), indicating that residues are compatible with their environments in the three-dimensional structure (Figure 2). There is, however, one segment of polypeptide chain, which includes Trp 227, that has an average profile score below 0.2 (Figure 2, broken line). This segment (loop B) makes packing contacts with a noncrystallographic symmetry-related molecule in the crystal. In particular, Trp 227 binds in an apolar cleft which influences its environment. Thus, when crystal contacts are included, the average profile score near residue 227 increases dramatically (Figure 2, solid line).
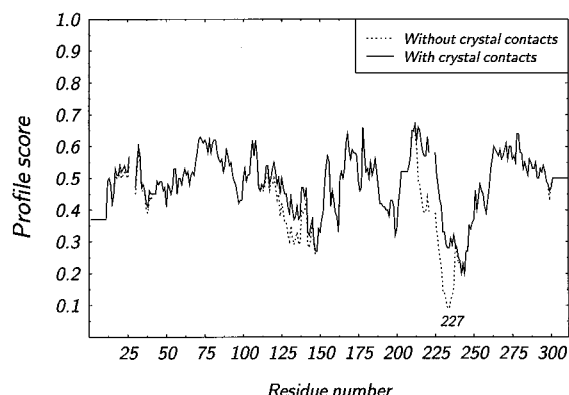
FIGURE 2: 3D−1D profile window plot. The 3D−1D profile score (Bowie et al., 1991) is averaged over a 21-residue window and plotted versus residue number for molecule 1 alone (broken lines) and in the presence of its noncrystallographic symmetry mate, molecule 2 (solid lines). In the crystal, the labeled segment containing Trp 227 is inserted into an apolar cleft in molecule 2. Residues 228−233 of one molecule and residues 131−135 of the other also form a crystal contact. Residues 27−29, 222−224, and 312−322 are disordered and were not included in the model.

Structure of the Cofactor Binding Site. Members of the SDR and AKR superfamilies use different structural motifs for binding cofactor. SDR superfamily members have a Rossmann nucleotide-binding fold (Rossmann et al., 1974; Jornvall et al., 1995) whereas AKR superfamily members bind cofactor through interactions with residues at the C-terminal end of the α/β barrel (Wilson et al., 1992; El-Kabbani et al., 1995; Wilson et al., 1995). Like other AKR superfamily members, 3α-HSD binds NADP$^+$ through interactions with residues in or near the C-terminal ends of several β-strands (β1, β2, β5, β6, β7, and β8) and an α-helix (α8) (Figures 1 and 3).

The electron density map shows clearly that NADP$^+$ is bound to 3α-HSD in two conformations that are rather similar (Figure 4). The only atoms that significantly differ between the two conformations are in the pyrophosphate bridge joining the bases. There are two distinct positions occupied by the adenosine 5'-phosphate and O5' atoms (Figure 4). The bases and sugars are identical in both conformations. The predominant conformation, with 65% occupancy (high occupancy conformation), is essentially the same as in the binary complexes of aldose reductase (Wilson et al., 1992), aldehyde reductase (El-Kabbani et al., 1995), and FR-1 (Wilson et al., 1995). The low (30%) occupancy conformation is unique and is probably permitted by the lack of salt links found in other AKR superfamily structures (the "safety belt", discussed below).

The orientation of NADP$^+$ in the cofactor binding site is determined by several hydrogen bonds and salt bridges (Figure 5). At the core of the barrel, the nicotinamide ring stacks against the side chain of Tyr 216 and is oriented by a network of three hydrogen bonds between the carboxamide group and the side chains of Ser 166, Asn 167, and Gln 190. These interactions would permit the 4-*pro-R* hydrogen in NADPH to point away from Tyr 216 and toward the active site (Figure 5a), consistent with the known stereochemistry of the reaction. These four side chains are identical or conservatively replaced among virtually all members of the AKR superfamily.

At the periphery of the barrel, the adenosine 2',5'-diphosphate portion of NADP$^+$ is bound by hydrogen bonds

and a salt bridge. In particular, several interactions between 3α-HSD and the 2'-phosphate may be the reason NADPH (K_d = 195 nM) binds more tightly than NADH (K_d = 160 μM). Ser 271 and Arg 276 form hydrogen bonds and a salt bridge with the 2'-phosphate in NADP$^+$ (see Figure 5b), and although the Arg 270−2'-phosphate salt bridge is unusually long, it may also provide some binding energy. These residues are highly conserved and perform similar roles in other AKR binary complex structures (Wilson et al., 1992; El-Kabbani et al., 1995; Wilson et al., 1995), suggesting the ability to bind NADP(H) with high affinity is important for enzyme function in the superfamily.

Comparison of the apoenzyme and binary complex structures reveals that several side chains undergo conformational changes upon NADP$^+$ binding (Figure 6). The largest changes occur near the adenine ring, where the side chain of Arg 276 is repositioned, moving its guanidinium group 5 Å and vacating the cavity where the adenine ring binds. In its new position, the side chain of Arg 276 forms a hydrogen bond and salt bridge with the 2'-phosphate of NADP$^+$. In order to accommodate the new position of Arg 276, Asn 273 adopts different side chain torsion angles χ_1 and χ_2. Ser 271 and Glu 279 also move to bind NADP$^+$ (Figure 6). These conformational changes could contribute to the rate-limiting binding of cofactor observed in the kinetic mechanism (Askonas et al., 1991). In contrast, side chains near the nicotinamide ring undergo only slight changes.

The Safety Belt in 3α-HSD. The 3α-HSD·NADP$^+$ binary complex structure differs from other AKR structures such as aldose reductase; in the latter structure, the pyrophosphate bridge of the cofactor lies in a tunnel formed by salt links between Asp 216 in loop B, Lys 21, and Lys 262. These salt links, referred to as a safety belt, lock the cofactor in the binding site and are formed or broken by a conformational change in loop B upon binding or release (Wilson et al., 1992; Borhani et al., 1992). Analogous salt links have been found in other AKR superfamily structures (El-Kabbani et al., 1995; Wilson et al., 1995).

The fact that binding and release of cofactor is rate limiting in the reaction catalyzed by 3α-HSD (Askonas et al., 1991) suggests that similar conformational changes in loop B may occur. However, sequence analysis reveals that two of the salt-linked residues in aldose reductase (Asp 216 and Lys 21) have no direct counterparts in 3α-HSD because of insertions in two loops, suggesting that 3α-HSD has an altered or absent safety belt (Wilson et al., 1995). An altered safety belt could conceivably be formed by a nearby lysine (Lys 28) and aspartic acid (Asp 224). Other HSDs in the AKR superfamily contain sequence insertions and may also have modified safety belts.

In the binary complex, crystal packing results in loop B (Trp 227) binding to an apolar cleft (discussed below), which does not allow the potential safety belt to form. Because of this packing interaction, it is possible that loop B does not adopt its normal solution conformation, which might involve a safety belt. However, the presence of this crystal contact could also indicate that the safety belt is weak or nonexistent; otherwise, loop B would not be free to move and the packing contact would not be formed.

Implications for Binding Steroid Hormones and Other Ligands. Only a crystal structure with bound substrate or inhibitor will provide an atomic model of how 3α-HSD recognizes apolar ligands. However, the binary complex
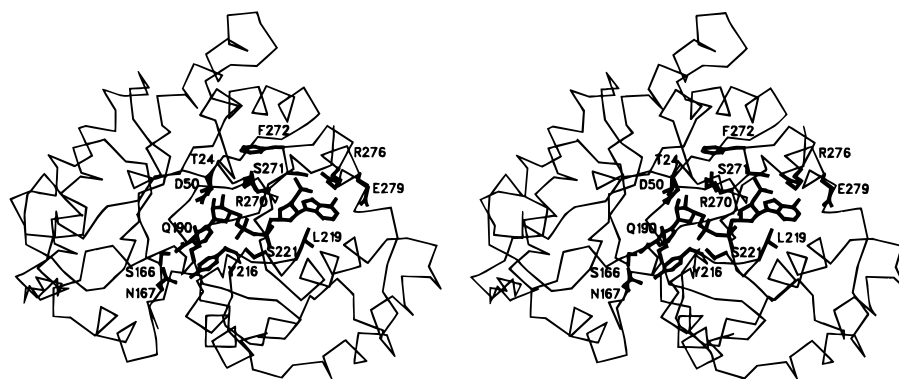
FIGURE 3: Location of the cofactor binding site. Stereo figures showing the Cα backbone of the protein and all non-hydrogen atoms of the high-occupancy conformation of NADP+. Side chains of residues that interact with both conformations of NADP+ are shown and labeled. Atomic interactions between 3α-HSD and NADP+ are shown in Figure 5.
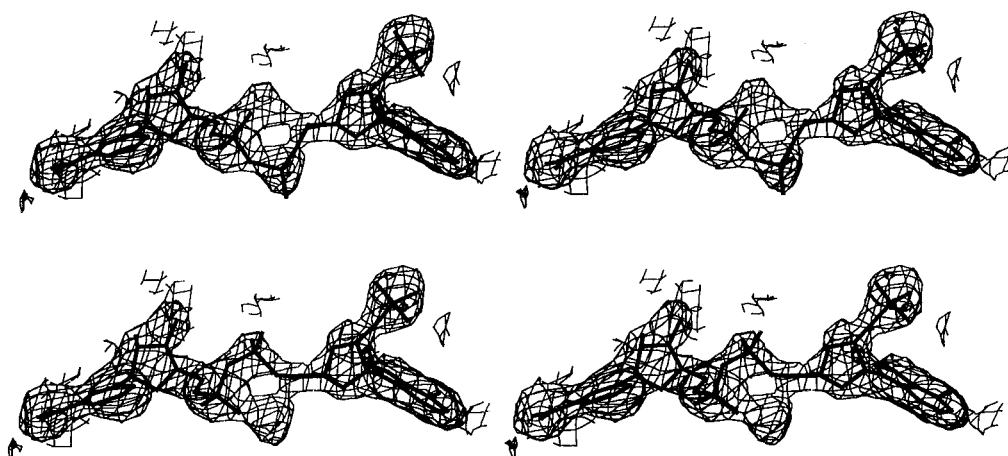

FIGURE 4: Electron density map, calculated with F_o amplitudes and averaged phases, contoured at 0.8σ around the two conformations of NADP+ in the active site. NADP+ coordinates were not used in calculating initial phases for averaging, so this region of the electron density is unbiased by the model. (a, top) The high-occupancy conformation (65% occupancy). (b, bottom) The low-occupancy conformation (30% occupancy). All map calculations used data in the resolution range 5−2.7 Å.

model allows us to identify the probable apolar substrate-binding cleft and to make several predictions about substrate recognition based on (1) a fortuitous crystal packing contact that may mimic a portion of a bound steroid hormone or other ligand, (2) an active site water molecule that is likely to correspond to the position of the substrate carbonyl or hydroxyl, and (3) site-directed mutagenesis data.

3α-HSD interacts with several types of apolar substrates of varying sizes and shapes, including hydroxysteroids, PAH *trans*-dihydrodiols, hydroxyprostaglandins, and aromatic ketones. These substrates are likely to bind to an apolar cleft formed by Leu 54, Trp 86, Leu 122, Phe 128, Phe 129, Leu 137, and Phe 139 on one side, Phe 118 and Met 120 on the bottom, and residues from the C-terminus and possibly Trp 227 with its associated loop on the opposing side. Because Trp 227 is involved in a crystal contact, the model does not provide its position in solution. However, mutations of Trp 86 and Trp 227 to tyrosine reduce steroid hormone binding up to 7-fold and 20-fold, respectively, suggesting that both residues are located in the steroid-binding cleft (Jez et al., 1995). This cleft could accommodate steroid hormones and other substrates and is proximal to the residues involved in catalysis, as well as the nicotinamide ring.

In this crystal form, part of the apolar cleft is used to create a crystal packing contact when loop B extends away from the surface of the other molecule in the asymmetric unit and Trp 227 of this neighboring molecule (Trp 227′) binds within

4 Å of eight residues in the apolar cleft (Figure 7). Although the electron density for portions of the flanking loop is diffuse, Trp 227′ is well ordered (Figure 7b). We suggest that Trp 227′ may mimic a portion of the distal end of a bound steroid hormone or other ligand on the basis of the following observations: (1) Trp 227′ interacts with Trp 86, which is implicated by site-directed mutagenesis in binding steroid hormones (Jez et al., 1995), (2) the distance between Trp 227′ and C4 of the nicotinamide ring is comparable to the length of a steroid nucleus or PAH *trans*-dihydrodiol (~11 Å), and (3) 92% of the solvent-accessible surface area of Trp 227′ is buried, as would be expected for a bound apolar ligand. Graphic modeling of steroid ligands based on the position of Trp 227′ and a bound water molecule (discussed below) shows that these ligands could fit into the apolar cleft but would require some accommodating structural changes.

The crystal contact provides insight into some of the conformational changes that may occur when 3α-HSD binds ligands. It might be expected that the C-terminus will be involved in substrate binding, analogous to residues 298−303 in aldose reductase (residues 306−311 in 3α-HSD), which undergo a conformational change upon binding the inhibitor zopolrestat (Wilson et al., 1993), and residues 303−315 (residues 311−322 in 3α-HSD), which are involved in substrate specificity and catalytic efficiency (Bohren et al., 1992). The binding of Trp 227′ in the apolar cleft similarly
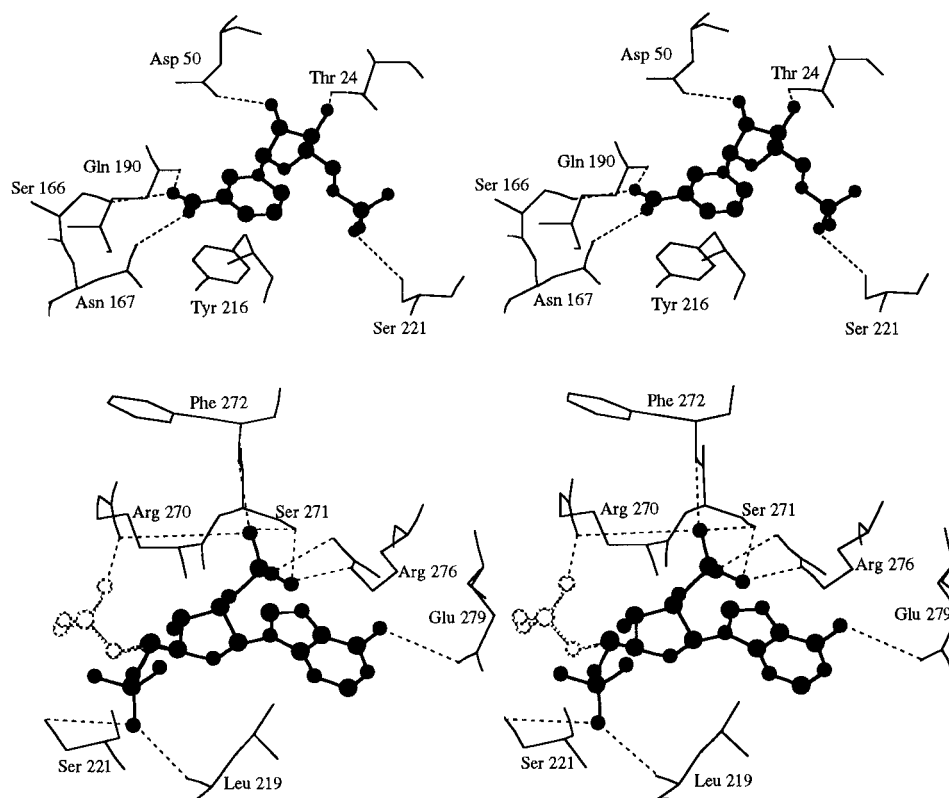
FIGURE 5: Atomic interactions binding NADP$^+$ to 3α-HSD. Stereo figures of the cofactor-binding site in approximately the same orientation as in Figure 3. The NADP$^+$ molecule is shown in ball-and-stick representation with atom sizes proportional to van der Waals radii. Hydrogen bonds and salt bridges are indicated by broken lines. (a, top) Nicotinamide ribose monophosphate portion of NADP$^+$ with the side chains of 3α-HSD to which it binds. (b, bottom) Adenosine 2′,5′-diphosphate portion of NADP$^+$ with the side chains of 3α-HSD to which it binds. The 5′- phosphate and O5′ atoms of the low-occupancy conformation of NADP$^+$ are shown in dotted lines. These are the only atoms whose positions significantly differ between the high- and low-occupancy conformations. This figure was prepared using MACIMDAD (Molecular Applications Group, Palo Alto, CA).
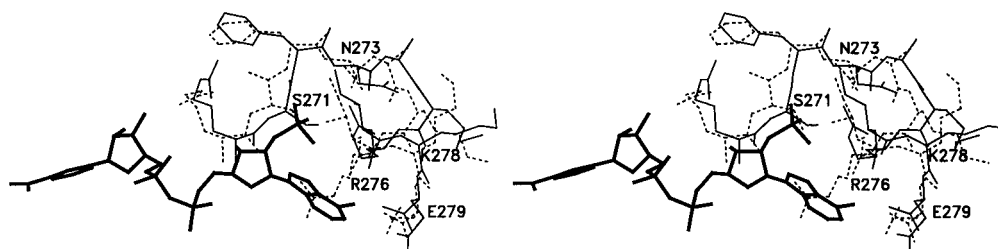

FIGURE 6: Side chain movements in the active site upon cofactor binding. Stereo figures showing all non-hydrogen atoms in residues 270−279 in apo 3α-HSD (Hoog et al., 1994) (broken lines) and in the 3α-HSD·NADP$^+$ binary complex (solid lines). All non-hydrogen atoms of the high-occupancy conformation of NADP$^+$ are shown in thick lines. Residues which undergo side chain movements with rms differences >2 Å are labeled.

involves some C-terminal residues in 3α-HSD. Residues 306−308 undergo a conformational change, and residues 309−311 are ordered in the binary complex structure, whereas they are disordered in the apoenzyme structure. In addition, Tyr 310 makes van der Waals contacts with Trp 227′, forming part of the apolar cleft (Figure 7). It is likely that the C-terminus also undergoes conformational changes to accommodate steroid hormones or other ligands and forms part of the apolar cleft.

A bound water molecule in the active site provides additional information about substrate binding. This water molecule forms hydrogen bonds with two of the amino acid residues implicated in binding and catalysis, His 117 and Tyr 55 (Figure 7a). In particular, Tyr 55, which may be the proton donor (discussed below), is 3.1 Å from this water molecule. Oxygen atoms in water and in inhibitors have been observed at this position in several crystallographic

structures of aldose reductase and its inhibitor complexes (Harrison et al., 1994; Wilson et al., 1992, 1993), and it seems likely that the oxygen atom in the carbonyl or hydroxyl group of a substrate will bind here as well. On the basis of this assumption and knowledge of the stereochemistry of hydride transfer, we propose that the α face of a bound steroid will be oriented toward the side of the apolar cleft that contains Trp 86.

Implications for the Catalytic Mechanism. The catalytic mechanism of 3α-HSD involves direct hydride transfer of the *pro-R* hydrogen from the C4 position of the nicotinamide ring to the acceptor carbonyl at C3 of the steroid substrate (Askonas et al., 1991). Since 3α-HSD is a nonmetallo-enzyme, the presence of a general acid to polarize the acceptor carbonyl is required to facilitate the reduction reaction. In the oxidation direction, the same amino acid may function as a general base. Identification of the amino
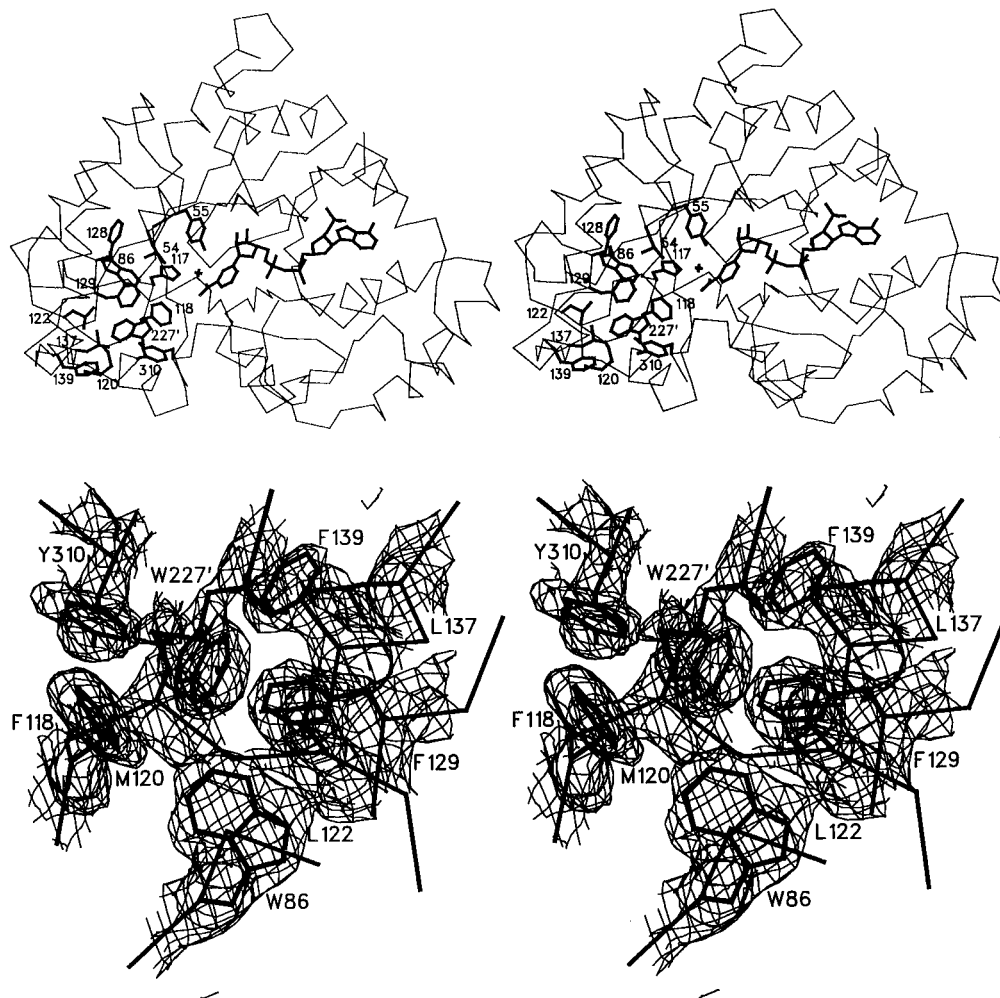
FIGURE 7: Crystal packing contact in the apolar cleft. Residues in the symmetry-related molecule are indicated with primed numbers. (a, top) Stereo figures of the crystal contact between Trp 227′ and the apolar cleft (top view). Trp 227′ and a bound water molecule may mimic portions of a bound ligand. Side chains in the apolar cleft (Leu 54, Trp 86, Phe 118, Met 120, Leu 122, Phe 128, Phe 129, Leu 137, Phe 139, Tyr 310) are shown, as well as Trp 227′ and two residues involved in substrate binding and catalysis (Tyr 55 and His 117). The high-occupancy conformation of NADP$^+$ is also shown. The active site water molecule is represented by its oxygen atom depicted as a cross. (b, bottom) Stereo figures of electron density, calculated with F_o amplitudes and averaged phases, contoured at 1σ around Trp 227′ and the residues in the apolar cleft with which it interacts at a distance of ≤ 4 Å.
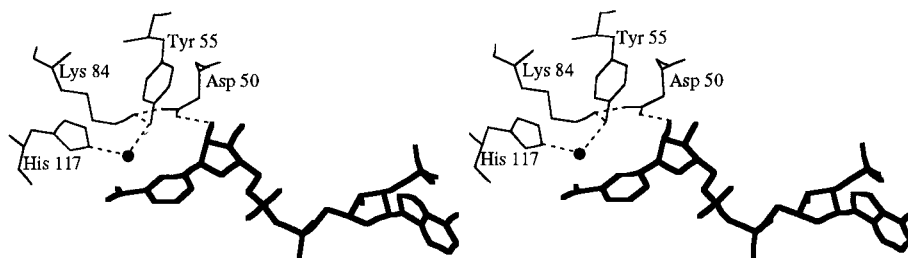

FIGURE 8: Residues near the nicotinamide ring. Stereo figures showing Asp 50, Tyr 55, Lys 84, His 117, and the high-occupancy conformation of NADP$^+$. The active site water molecule is represented by its oxygen atom. Hydrogen bonds are shown as broken lines. This figure was prepared using MACIMDAD (Molecular Applications Group, Palo Alto, CA).

acid residue which functions as the general acid is an important step toward understanding the catalytic mechanism of this enzyme.

A catalytic mechanism has been proposed for 3α-HSD and other enzymes in the AKR superfamily which involves three residues near the nicotinamide ring (Figure 8). In this mechanism, a tyrosine (Tyr 55 in 3α-HSD) acts as the general acid, a lysine (Lys 84 in 3α-HSD) decreases the pK_a of the tyrosine by hydrogen bonding to it, and an aspartic acid (Asp 50 in 3α-HSD) is salt-bridged to the lysine (Bohren et al., 1994; Hoog et al., 1994). A histidine (His 117 in

3α-HSD) has also been implicated in catalysis in aldose reductase on the basis of its position and pK_a, which is near the pH optimum of the reaction (Liu et al., 1993).

The catalytic mechanism of 3α-HSD has been investigated by site-directed mutagenesis of these four residues. These experiments show that Asp 50 and His 117 mutants have low residual activity for steroid oxidation (B. P. Schlegel, manuscript in preparation). In contrast, a Y55F mutant was found to be inactive using steroid substrates, while retaining the ability to form E·NADPH and E·NADH·testosterone complexes with the K_d for the binding cofactor unchanged.
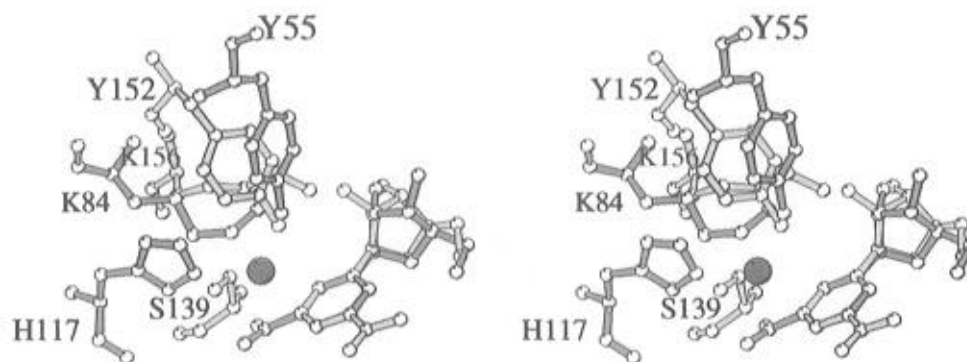
FIGURE 9: Comparison of AKR and SDR enzymes in which the protein folds differ but the active site residues superimpose. Stereo figures of active site residues in 3α-HSD (blue) superimposed with those in 3α,20β-HSD, a member of the SDR superfamily (yellow) (Ghosh et al., 1994). All non-hydrogen atoms in Tyr 55, Lys 84, His 117, and nicotinamide ribose in the 3α-HSD·NADP⁺ binary complex and in Ser 139, Tyr 152, Lys 156, and nicotinamide ribose in the 3α,20β-HSD·NAD⁺ binary complex are shown. The water molecule in 3α-HSD that may mimic the carbonyl oxygen in a 3-ketosteroid substrate is represented by its oxygen atom in red. The superposition was based on the nicotinamide ring position, excluding the carboxamide substituent at the C3 position. The reason for excluding this substituent was that 3α,20β-HSD transfers the *pro-S* hydrogen, while 3α-HSD transfers the *pro-R* hydrogen, so that although the nicotinamide rings lie in the same plane, they are flipped 180° relative to one another and the carboxamide groups do not superimpose. This figure was prepared using MOLSCRIPT (Kraulis, 1991).

Although the affinity for steroid was reduced over 30-fold (Pawlowski & Penning, 1994), this decrease is insufficient to explain the complete loss of catalytic activity, suggesting that Tyr 55 could be the general acid. Similarly, Lys 84 mutants are inactive, suggesting this residue is involved in the catalytic mechanism (B. P. Schlegel, manuscript in preparation).

The binary complex model and the inactive Tyr 55 and Lys 84 mutants support the proposed catalytic mechanism in which Tyr 55 is the general acid and in which this function may be facilitated by Lys 84. Tyr 55 is close to the C4 position of the nicotinamide ring and forms a hydrogen bond with the water molecule that may indicate the position of the carbonyl oxygen in a 3-ketosteroid substrate (Figure 8). In order for Lys 84 to interact directly with the water molecule (and hence the substrate carbonyl), the enzyme would have to undergo conformational changes involving Lys 84, His 117, and Tyr 55. On this basis, it is less likely that Lys 84 would act as the general acid, leaving Tyr 55 as the most viable candidate. It should be noted that His 117 also interacts with the water molecule and is in a reasonable position to participate in catalysis, although the enzyme retains low activity upon its mutation.

Comparison of AKR and SDR Enzymes: Superposition of Active Site Residues Suggests Convergent Evolution to a Common Catalytic Mechanism. In addition to the AKR superfamily, the structurally distinct SDR superfamily (Jornvall, 1995) includes several HSDs. Enzymes in the AKR and SDR families share no significant sequence identity or structural homology, although they interconvert similar, and in some cases identical, substrates. Known structures in the SDR superfamily include human placental 17β-HSD (Ghosh et al., 1995), bacterial 3α,20β-HSD (Ghosh et al., 1994), and rat liver dihydropteridine reductase (Varughese et al., 1992).

Within the SDR superfamily, members share a consensus active site sequence Tyr-X-X-X-Lys. The invariant tyrosine and lysine residues are essential for catalysis (Ensor & Tai, 1991; Chen et al., 1993; Obeid & White, 1992) and are located in the active site (Ghosh et al., 1994, 1995; Varughese et al., 1992). Prior to the structure determination of 3α-

HSD, it was noted that its sequence contains this motif at residues 205−209 (Tyr-Cys-Lys-Ser-Lys), and it was hypothesized that Tyr 205 was the general acid. However, the apoenzyme structure revealed that the consensus sequence is located in α-helix 6 (Hoog et al., 1994) and is distant from the active site. This observation suggested Tyr 205 was not involved in catalysis by 3α-HSD, which was later confirmed by site-directed mutagenesis (Pawlowski & Penning, 1994).

Although the Tyr-X-X-X-Lys sequence is not essential for catalysis in 3α-HSD, there is a tyrosine/lysine pair near NADP⁺ in the active site of the binary complex (Tyr 55 and Lys 84; see Figure 8). Like the SDR tyrosine/lysine pair, these residues are virtually invariant in the AKR superfamily and have been implicated in catalysis. The fact that SDR and AKR HSDs interconvert similar substrates, using the same conserved amino acid side chains suggested that they might have similar catalytic mechanisms although their primary sequences and protein folds differ. To explore this hypothesis, we superimposed the active sites of 3α-HSD and the SDR enzyme 3α,20β-HSD (Ghosh et al., 1994) on the basis of the nicotinamide ring positions and examined the positions of active site residues (Figure 9). Remarkably, the oxygen atoms in the tyrosine hydroxyl groups superimpose to within 0.5 Å. The lysine residues also adopt similar positions with their amino groups separated by 2.2 Å. In addition, the Nε2 atom of His 117 in 3α-HSD and the hydroxyl group of Ser 139 in 3α,20β-HSD are separated by only 1.7 Å and are both within hydrogen-bonding distance from the water molecule in 3α-HSD that may indicate the position of the substrate carbonyl.

Thus, although AKR and SDR superfamily enzymes have completely different protein folds, the conserved tyrosine/lysine pair is present in similar positions in both. In addition, His 117 and Ser 139 are present in similar positions and are highly conserved in their respective superfamilies. Our observation of the superposition of the active site residues in the two superfamilies suggests convergent evolution of a common reaction mechanism, involving a lysine and a tyrosine which may act as the general acid.

ACKNOWLEDGMENT

REFERENCES

Askonas, L. J., Ricigliano, J. W., & Penning, T. M. (1991) *Biochem. J. 278*, 835−841.

Bohren, K. M., Grimshaw, C. E., & Gabbay, K. H. (1992) *J. Biol. Chem. 267*, 20965−20970.

Bohren, K. M., Grimshaw, C. E., Lai, C.-J., Harrison, D. H., Ringe, D., Petsko, G. A., & Gabbay, K. H. (1994) *Biochemistry 33*, 2021−2032.

Borhani, D. W., Harter, T. M., & Petrash, J. M. (1992) *J. Biol. Chem. 267*, 24841−24847.

Bowie, J. U., Luthy, R., & Eisenberg, D. (1991) *Science 253*, 164−170.

Bruce, N. C., Willey, D. L., Coulson, A. F. W., & Jeffery, J. (1994) *Biochem. J. 299*, 805−811.

Brunger, A. T. (1992a) *X-PLOR Version 3.1: A system for X-ray crystallography and NMR*, Yale University Press, New Haven, CT.

Brunger, A. T. (1992b) *Nature 355*, 472−475.

Chen, Z., Jiang, J. C., Lin, Z.-G., Lee, W. R., Baker, M. E., & Chang, S. H. (1993) *Biochemistry 32*, 3342−3346.

Collaborative Computational Crystallography Project, Number 4 (1994) *Acta Crystallogr. D 50*, 760−763.

Deyashiki, Y., Ogasawara, A., Nakayama, T., Nakanishi, M., Miyabe, Y., Sato, K. & Hara, A. (1994) *Biochem. J. 299*, 545−552.

Deyashiki, Y., Ohshima, K., Nakanishi, M., Sato, K., Matsuura, K., & Hara, A. (1995) *J. Biol. Chem. 270*, 10461−10467.

Dufort, I., Soucy, P., Zhang, Y., & Luu-The, V. (1995) *Cloning and characterization of human type 2 3α-hydroxysteroid dehydrogenase from human prostatic cDNA library*, Fifth International Congress, Hormones and Cancer, Quebec City, Canada.

El-Kabbani, O., Judge, K., Ginell, S. L., Myles, D. A. A., DeLucas, L. J., & Flynn, T.G. (1995) *Nat. Struct. Biol. 2*, 687−692.

Ensor, C. M., & Tai, H.-H. (1991) *Biochem. Biophys. Res. Commun. 176*, 840−845.

Evans, S. V. (1993) *J. Mol. Graphics 11*, 134−138.

Furey, W., & Swaminathan, S. (1990) *PHASES—a program package for the processing and analysis of diffraction data for macromolecules*, American Crystallographic Association Meeting, New Orleans, LA.

Ghosh, D., Wawrzak, Z., Weeks, C. M., Duax, W. L., & Erman, M. (1994) *Structure 2*, 629−640.

Ghosh, D., Pletnev, V. Z., Zhu, D.-W., Wawrzak, Z., Duax, W. L., Pangborn, W., Labrie, F., & Lin, S.-X. (1995) *Structure 3*, 503−513.

Harrison, D. H., Bohren, K. M., Ringe, D., Petsko, G. A., & Gabbay, K. H. (1994) *Biochemistry 33*, 2011−2020.

Hoog, S. S., Pawlowski, J. E., Alzari, P. M., Penning, T. M., & Lewis, M. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 2517−2521.

Jez, J. M., & Penning, T. M. (1995) *The Role of Tryptophans in Ligand Recognition by Rat Liver 3α-Hydroxysteroid Dehydrogenase*, International Symposium on DHEA Transformation into Androgens and Estrogens in Target Tissues, Quebec City, Canada.

Jones, T. A., Zou, J.-Y., Cowan, S. W., & Kjeldgaard, M. (1991) *Acta Crystallogr. A 47*, 110−119.

Jornvall, H., Persson, B., Krook, M., Atrian, S., Gonzalez-Duarte, R., Jeffery, J., & Ghosh, D. (1995) *Biochemistry 34*, 6003−6013.

Kabsch, W. (1988) *J. Appl. Crystallogr. 21*, 916−924.

Khanna, M., Qin, K.-N., Wang, R. W., & Cheng, K.-C. (1995) *J. Biol. Chem. 270*, 20162−20168.

Kraulis, P. (1991) *J. Appl. Crystallogr. 24*, 946−950.

Lacy, W. R., Washenick, K. J., Cook, R. G., & Dunbar, B. S. (1993) *Mol. Endocrinol. 7*, 58−66.

Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.

Leslie, A. G. W. (1992) *CCP4 and ESF-EAC MB Newsletter on Protein Crystallography*, Vol. 26.

Lin, H.-K., Jez, J. M., & Penning, T. M. (1995) *Cloning of a human prostate cDNA with high sequence identity to rat liver 3α-hydroxysteroid/dihydrodiol dehydrogenase*, International Symposium on DHEA Transformation into Androgens and Estrogens in Target Tissues, Quebec City, Canada.

Liu, S.-Q., Bhatnagar, A., & Srivastava, S. K. (1993) *J. Biol. Chem. 268*, 25494−25499.

Matthews, B. W. (1968) *J. Mol. Biol. 33*, 491−497.

Miura, R., Shiota, K., Noda, K., Yagi, S., Ogawa, T., & Takahashi, M. (1994) *Biochem. J. 299*, 561−567.

Mukharji, I. (1982) Ph.D. Dissertation, The Johns Hopkins University, Baltimore, MD.

Obeid, J., & White, P. C. (1992) *Biochem. Biophys. Res. Commun. 188*, 222−227.

Pawlowski, J. E., & Penning, T. M. (1994) *J. Biol. Chem. 269*, 13502−13510.

Pawlowski, J. E., Huizinga, M., & Penning, T. M. (1991) *J. Biol. Chem. 266*, 8820−8825.

Penning, T. M. (1993) *Chem.-Biol. Interact. 89*, 1−34.

Penning, T. M., Mukharji, I., Barrows, S., & Talalay, P. (1984) *Biochem. J. 222*, 601−611.

Penning, T. M., Ohnishi, T., Ohnishi, T. S., & Harvey, R. G. (1996) *Chem. Res. Toxicol. 9*, 84−92.

Rondeau, J. M., Tete-Favier, F., Podjarny, A, Reymann, J. M., Barth, P., Biellmann, J. F., & Moras, D. (1992) *Nature 355*, 469−472.

Rossmann, M. G., Moras, D., & Olsen, K. W. (1974) *Nature 250*, 194−199.

Smithgall, T. E., Harvey, R. G., & Penning, T. M. (1988) *J. Biol. Chem. 263*, 1814−1820.

Stolz, A., Hammond, L., Lou, H., Takikawa, H., Ronk, M., & Shively, J. E. (1993) *J. Biol. Chem. 268*, 10448−10457.

Varughese, K. I., Skinner, M. M., Whiteley, J. M., Matthews, D. A., & Xuong, N. H. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 6080−6084.

Wilson, D. K., Bohren, K. M., Gabbay, K. H., & Quiocho, F. A. (1992) *Science 257*, 81−84.

Wilson, D. K., Tarle, I., Petrash, J. M., & Quiocho, F. A. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 9847−9851.

Wilson, D. K., Nakano, T., Petrash, J. M., & Quiocho, F. A. (1995) *Biochemistry 34*, 14323−14330.