REVIEW ARTICLE

Comparative anatomy of the aldo–keto reductase superfamily

Joseph M. JEZ*, Melanie J. BENNETT*, Brian P. SCHLEGEL†, Mitchell LEWIS* and Trevor M. PENNING†‡

Departments of *Biochemistry & Biophysics and †Pharmacology, University of Pennsylvania Medical School, 3620 Hamilton Walk, Philadelphia, PA 19104, U.S.A.

The aldo–keto reductases metabolize a wide range of substrates and are potential drug targets. This protein superfamily includes aldose reductases, aldehyde reductases, hydroxysteroid dehydrogenases and dihydrodiol dehydrogenases. By combining multiple sequence alignments with known three-dimensional structures and the results of site-directed mutagenesis studies, we have developed a structure/function analysis of this superfamily. Our studies suggest that the $(\alpha/\beta)_8$-barrel fold provides a common scaffold for an NAD(P)(H)-dependent catalytic activity, with substrate specificity determined by variation of loops on the C-terminal side of the barrel. All the aldo–keto reductases are dependent on nicotinamide cofactors for catalysis and retain a similar cofactor binding site, even among proteins with less than 30% amino acid sequence identity. Likewise, the aldo–keto reductase active site is highly conserved. However, our alignments indicate that variation of a single residue in the active site may alter the reaction mechanism from carbonyl oxidoreduction to carbon–carbon double-bond reduction, as in the 3-oxo-5β-steroid 4-dehydrogenases (Δ^4-3-ketosteroid 5β-reductases) of the superfamily. Comparison of the proposed substrate binding pocket suggests residues 54 and 118, near the active site, as possible discriminators between sugar and steroid substrates. In addition, sequence alignment and subsequent homology modelling of mouse liver 17β-hydroxysteroid dehydrogenase and rat ovary 20α-hydroxysteroid dehydrogenase indicate that three loops on the C-terminal side of the barrel play potential roles in determining the positional and stereo-specificity of the hydroxysteroid dehydrogenases. Finally, we propose that the aldo–keto reductase superfamily may represent an example of divergent evolution from an ancestral multifunctional oxidoreductase and an example of convergent evolution to the same active-site constellation as the short-chain dehydrogenase/reductase superfamily.

INTRODUCTION

Traditionally, oxidoreductases belong to either the long-chain alcohol dehydrogenase or the short-chain dehydrogenase/reductase (SDR; formerly the short-chain alcohol dehydrogenase) superfamilies. Yeast and liver alcohol dehydrogenases typify the long-chain alcohol dehydrogenases, which use zinc for catalysis [1]. The SDRs are multimeric non-metallo-oxidoreductases with a conserved Tyr-Xaa-Xaa-Xaa-Lys catalytic motif and a Rossmann fold for NAD(P)(H) binding [2]. These two superfamilies have dissimilar three-dimensional structures and are reliant on either a zinc atom or a catalytic acid/base to perform their reactions [3,4]. An increasing number of oxidoreductases not related to either of these superfamilies have been identified by cDNA cloning; these enzymes are similar to each other, and belong to the aldo–keto reductase (AKR) superfamily [5,6].

The AKRs are monomeric proteins, about 320 residues in size, that bind nicotinamide cofactor without a Rossmann-fold motif [7–10]. Found in mammals, amphibians, plants, yeast, protozoa and bacteria, these proteins metabolize a diverse range of substrates, including aliphatic and aromatic aldehydes, monosaccharides, steroids, prostaglandins, polycyclic aromatic hydrocarbons and isoflavinoids [6,11–13].

Many of the mammalian AKRs are potential therapeutic targets, and structure-based drug design may lead to compounds with the desired specificity and clinical efficacy. Aldose reductase (ADR; EC 1.1.1.21) converts circulating glucose into sorbitol, a hyperosmotic sugar, and may play a role in diabetic retinopathy, neuropathy and nephropathy. As such, ADR inhibitors could prove useful as treatments for these complications of diabetes [14]. The clinically tested ADR inhibitors are from two main chemical classes: aryl-substituted acetic acids (ponalrestat and zopolrestat) and acidic cyclic imides, particulary spiro-hydantoins (sorbinil) and spiro-succinimides (alrestatin) [15]. The clinical potential of designed inhibitors has driven the investigation of the structure and function of ADR. Similarly, the mammalian aldehyde reductases (ALRs; EC 1.1.1.2) play a significant role in the metabolism of neurotransmitter aldehydes produced by monoamine oxidase [16], and ALR inhibitors may have anti-depressant properties.

Both 3α-hydroxysteroid dehydrogenase (3α-HSD; EC 1.1.1.213; formerly EC 1.1.1.50, but renamed due to A-face specific hydride transfer) and prostaglandin F synthase (EC 1.1.1.188) are targets for non-steroidal anti-inflammatory drugs, and may regulate levels of inflammatory prostaglandins [17,18]. Other hydroxyprostaglandin dehydrogenases, involved in the metabolism of bronchoconstrictors, that may yet be assigned to this superfamily could be targets in the treatment of asthma. In addition, the HSDs of this superfamily have the potential to act as molecular switches, by converting potent steroid hormones into inactive metabolites, thereby regulating the amount of

Abbreviations used: SDR, short-chain dehydrogenase/reductase (formerly short-chain alcohol dehydrogenase); AKR, aldo–keto reductase; ADR, aldose reductase (sorbitol-6-phosphate dehydrogenase/xylose reductase; EC 1.1.1.21); ALR, aldehyde reductase (EC 1.1.1.2); HSD, hydroxysteroid dehydrogenase; DD, dihydrodiol dehydrogenase (EC 1.3.1.20); FR-1, fibroblast growth factor-1-induced protein.

‡ To whom correspondence should be addressed.

Table 1 Members of the AKR superfamily

The list of AKRs was compiled by screening the GenBank Nucleotide Sequence Database, the Protein Identification Resource Database and the SwissProt Database with the BLAST program [23,24] using the amino acid sequence of rat liver 3α-HSD as a search query [25]. The AKR superfamily nomenclature indicated in parentheses has been described elsewhere [26].

Designation	Species	Enzyme	Reference
Hum-ALR	*Homo sapiens* (liver)	ALR (AKR1A1)	42
Por-ALR	*Sus scrofa*	ALR (AKR1A2)	43
Rat-ALR	*Rattus norvegicus* (liver)	ALR (AKR1A3)	44
Mou-LKR	*Mus musculus* (liver)	Ketoreductase (AKR1B10)	41
Hum-ADR	*Homo sapiens* (placenta)	ADR (AKR1B1)	5, 34
Rab-ADR	*Oryctolagus cuniculus*	ADR (AKR1B2)	35
Mou-ADR	*Mus musculus* (kidney)	ADR (AKR1B3)	36
Rat-ADR	*Rattus norvegicus* (lens)	ADR (AKR1B4)	37
Bov-ADR	*Bos taurus* (lens/testis)	ADR (AKR1B5)	38, 39
Por-ADR	*Sus scrofa* (lens)	ADR (AKR1B6)	40
Mou-VDP	*Mus musculus*	Androgen-dependent vas deferens protein (AKR1B7)	58
Mou-FR1	*Mus musculus*	FR-1 (AKR1B8)	65
CHO-AKR	Chinese hamster ovary cells	Tripeptide-inducible AKR from CHO cells (AKR1B9)	64
Hum-20aHSD	*Homo sapiens* (liver)	20α-HSD (AKR1C1) (also DD1)	29, 30
Hum-BABP	*Homo sapiens* (liver)	Bile acid binding protein (AKR1C2) (also DD2 and 3α-HSD type III)	29, 31
Hum-3aHSDII	*Homo sapiens* (liver)	3α-HSD type II (AKR1C3)	32
Hum-ChlorRed	*Homo sapiens* (liver)	Chlordecone reductase (AKR1C4) (also DD4 and 3α-HSD type I)	31–33
Rab-20aHSD	*Oryctolagus cuniculus* (ovary)	20α-HSD (AKR1C5)	20
Mou-17bHSD	*Mus musculus* (liver)	17β-HSD (AKR1C6)	27
Bov-Pgs	*Bos taurus* (lung)	Prostaglandin F synthase (AKR1C7)	18
Bov-DD3	*Bos taurus* (liver)	DD3 (AKR1C11)	56
Rat-20aHSD	*Rattus norvegicus* (ovary)	20α-HSD (AKR1C8)	28
Rat-3aHSD	*Rattus norvegicus* (liver)	3α-HSD (AKR1C9) (also DD)	25
Frog-Rho	*Rana catesbeiana*	ρ-Crystallin (AKR1C10)	57
Hum-3o5bred	*Homo sapiens* (liver)	3-Oxo-5β-steroid 4-dehydrogenase (AKR1D1)	48
Rat-3o5bred	*Rattus norvegicus* (liver)	3-Oxo-5β-steroid 4-dehydrogenase (AKR1D2)	49
Bar-ALR	*Hordeum vulgare*	ALR (AKR1E1)	45
Brgr-ALR	*Bromus inermis*	ALR (AKR1E2)	46
App-S6Pdh	*Malus domestica*	Sorbitol phosphate dehydrogenase (AKR2A)	59
Pic-XylRed	*Pichia stipitis*	Xylose reductase (AKR2B1)	60
Klu-XylRed	*Kluyveromyces lactis*	Xylose reductase (AKR2B2)	61
Muc-TADH	*Mucor mucedo*	4-Dihydromethyltrisporate dehydrogenase (AKR2C)	62
Spor-ALR	*Sporidiobolus salmonicolor*	ALR (AKR3A)	47
Sac-GCY	*Saccharomyces cerevisiae*	GCY protein (AKR3B)	63
Alf-ChalRed	*Medicago sativa*	Chalcone reductase (AKR4A1)	50
Soy-ChalRed	*Glycine max*	Chalcone reductase (AKR4A2)	12
Gly-PKR	*Glycyrrhiza echinata*	Polyketide reductase (AKR4A3)	51
Lei-putRed	*Leishmania major*	Putative reductase (AKR5A)	55
Ps-Mordh	*Pseudomonas putida*	Morphine dehydrogenase (AKR5B)	54
Cb-25dkg	*Corynebacterium* sp.	2,5-Diketo-D-gulonate reductase (AKR5C)	52
Cb2-25dkg	*Corynebacterium* sp.	2,5-Diketo-D-gulonate reductase (AKR5D)	53
Rat-AFBred	*Rattus norvegicus* (liver)	Aflatoxin B1 ALR (AKR7)	66

hormone that can bind and activate nuclear receptors [19,20]. This role has been well documented for type 2 11β-HSD (EC 1.1.1.146) and type 1 17β-HSD (EC 1.1.1.62), which regulate the occupancy of mineralocorticoid and oestrogen receptors respectively. Although these HSDs belong to the SDR superfamily, HSDs of the AKR superfamily could perform a similar function. Thus therapeutics directed against tissue-specific HSDs may modulate occupancy of steroid-hormone receptors.

Zopolrestat and ponalrestat bind to ADR with nanomolar affinity, but are weak inhibitors of 3α-HSD, a member of the AKR superfamily [21]. This suggests that the design of inhibitors that are specific for members of the superfamily can be achieved. However, there has been little further progress towards this goal. Before targeting individual AKRs for rational drug design, an understanding of the relationship between structure and function in the superfamily is required.

In this review, we demonstrate the high degree of amino acid sequence similarity among the AKRs, and address how studies of 3α-HSD, ADR and ALR provide insight into the structure/function relationships across the superfamily. Amino acid comparisons suggest that proteins of the superfamily share a common $(\alpha/\beta)_8$-barrel three-dimensional fold and have a highly conserved nicotinamide-cofactor-binding pocket. While the AKR active site is also conserved both structurally and in sequence, our alignments suggest that a single amino acid change in this site may alter catalytic activity from carbonyl oxidoreduction to carbon–carbon double-bond reduction. In addition, our work indicates that the three loops on the C-terminal side of the structure are vital in discriminating between different substrates of the AKRs, especially for the HSDs of the superfamily. Finally, we raise the possibility of functional divergence within the superfamily from a common ancestral oxidoreductase, and of convergent evolution to the same catalytic mechanism as the SDR superfamily. The concept that a common ancestral oxidoreductase has undergone divergent evolution in the SDR superfamily with convergence to a common reaction mechanism (e.g. type I–IV 17β-HSD) has been raised previously [22]. Since both AKRs and SDRs contain alternating arrangements of α-helix

and β-sheet, albeit with different structural motifs, these proteins may have arisen from a now extinct common structural ancestor [2].

THE AKR SUPERFAMILY

A total of 42 oxidoreductases have been identified as members of the AKR superfamily (Table 1), including seven HSDs [five of which are also dihydrodiol dehydrogenases (DDs; EC 1.3.1.20)] [20,25,27–33], seven ADRs [5,34–41], six ALRs [5,42–47], two 3-oxo-5β-steroid 4-dehydrogenases (Δ^4-3-ketosteroid 5β-reductases) (ED 1.3.99.6) [48,49], three plant chalcone reductases [12,50,51], two 2,5-diketo-gulonate reductases from *Corynebacterium* [52,53], *Pseudomonas* morphine dehydrogenase (EC 1.1.1.218) [54], a putative reductase from *Leishmania* [55], human liver bile acid binding protein [29,31], bovine lung prostaglandin F synthase [17], bovine DD3 [56], frog lens ρ-crystallin [57], a mouse vas deferens protein [58], apple sorbitol-6-phosphate reductase [59], two xylose reductases from yeast [60,61], 4-dihydromethyltrisporate dehydrogenase from *Mucor mucedo* [62], yeast GCY protein [63], an inducible protein from Chinese hamster ovary cells [64], a murine protein induced by fibroblast growth factor-1 (FR-1) [65], and an ethoxyquin-inducible aflatoxin B1 aldehyde reductase [66]. It should be noted that the β-subunits of rat and bovine *Shaker* K^+ channels may be distantly related to the AKRs, but were not included in this analysis, since they do not exhibit any known catalytic activity [67].

Cluster analysis of the AKRs indicates that these proteins form six distinct groups within the superfamily (Figure 1). The most studied, and largest, group consists of the plant ALRs and the mammalian HSDs, ALRs and ADRs. Other families include proteins from yeast, prokaryotes and protozoa that catalyse reactions with substrates similar to those of ADR and ALR, but these proteins differ significantly from the latter enzymes in amino acid sequence. Three groups of the superfamily include proteins produced as a response to exogenous factors. For example, among the ADR family, FR-1 is induced by fibroblast growth factor [65], the mouse vas deferens protein displays

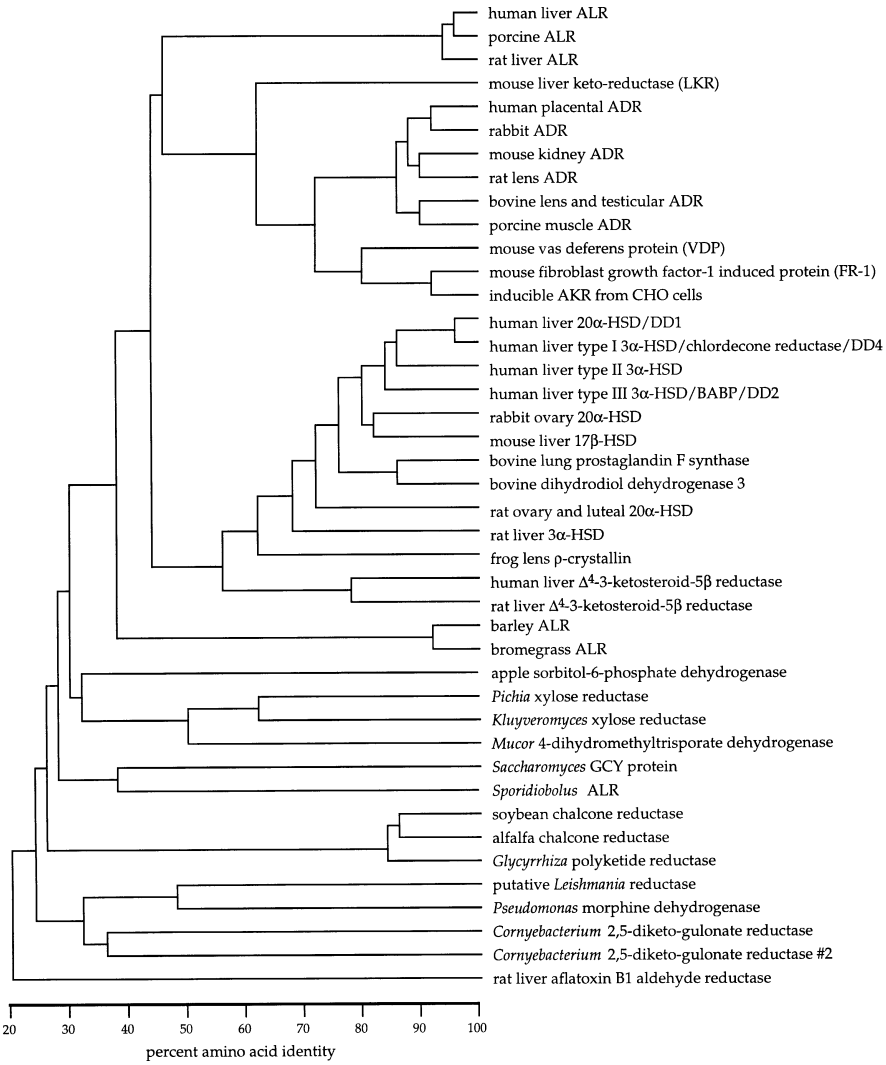


Figure 1 Cluster analysis of the AKR superfamily

The dendrogram is based on pairwise sequence alignments, and indicates the relative percentage amino acid identity among proteins of the AKR superfamily. Abbreviation: BABP, bile acid binding protein.

androgen-dependent expression [58], and an inducible AKR from Chinese hamster ovary cells has recently been cloned [64]. In addition, the plant chalcone reductases synthesize isoflavinoid phytoalexins in conjunction with chalcone synthase as an induced response to pathogen attack [12], while metabolism of the dialdehyde of aflatoxin B1 occurs through an ethoxyquin-inducible ALR [66].

A COMMON THREE-DIMENSIONAL STRUCTURE

Since members of the AKR superfamily show amino acid sequence identity, it is likely that they retain a common three-dimensional fold. The structures of human placental and bovine muscle ADRs, rat liver 3α-HSD, pig muscle ALR and murine FR-1 adopt an identical $(\alpha/\beta)_8$-barrel fold, as determined by X-ray crystallography [7–10,69]. These structures have eight parallel β-strands, and each β-strand alternates with an α-helix running anti-parallel to the strand, thereby forming the staves of the barrel (Figure 2).

Strict conservation of amino acid sequence occurs at 11 positions in the primary structure of all the AKRs: Gly-22, Gly-45, Asp-50, Lys-84, Asp-112, Pro-119, Gly-164, Asn-167, Pro-186, Gln-190 and Ser-271 (Figure 3). (The numbering of amino acid residues throughout this review corresponds to that of rat liver 3α-HSD.) Also, an additional eight positions are conserved in 41 of the AKRs: Gly-20, Tyr-55, Gly-62, Leu-113, Trp-148, Gly-158, Glu-192 and Arg-276. Five of these residues function in cofactor binding (Asp-50, Asn-167, Gln-190, Ser-271 and Arg-276), and three form parts of the active site (Asp-50, Tyr-55 and Lys-84). The remaining residues may play structural roles in forming the barrel core, since they are found within the β-strands, α-helices and short loops of the barrel. This pattern of conservation is not unusual. In other $(\alpha/\beta)_8$-barrel proteins, recurrent short loop structures between the α-helices and the β-strands of the barrel are thought to maintain the general size and ellipticity of the barrel [70–72].

Since the AKRs consist of approximately the same number of residues, we weighted the multiple alignments against gap insertion. The resulting multiple sequence alignment contained no major gaps in regions that correspond to the components of the barrel structure, further supporting a conserved three-dimensional fold for the superfamily. However, major gaps in the alignments appear in the sequences of three large loops on the C-terminal side of the barrel (Figures 2 and 3). The AKRs seemingly maintain the architecture of the barrel scaffold while tailoring substrate specificity through modification of the loops near the active site, as discussed below.

KINETIC AND CATALYTIC MECHANISM

The AKRs catalyse oxidation and reduction reactions on a range of substrates using NAD(P)(H) as cofactor. An ordered Bi Bi kinetic mechanism, in which cofactor binds first and leaves last, has been demonstrated for ALR, ADR, 3α-HSD and 3-oxo-5β-steroid 4-dehydrogenase, and may be a hallmark feature of other AKRs [73–76]. The reduction reaction involves 4-*pro-R* hydride transfer from NAD(P)H to the substrate carbonyl and protonation of the oxygen by a residue of the enzyme acting as a general acid. The reverse occurs in the oxidation reaction. Whether this reaction has a concerted mechanism or two distinct steps, with either hydride or proton transfer occurring first, remains unclear. As discussed below, the 5β-reductases of the superfamily represent a variation on this general mechanism, and reduce a carbon–carbon double bond instead of a carbonyl group.

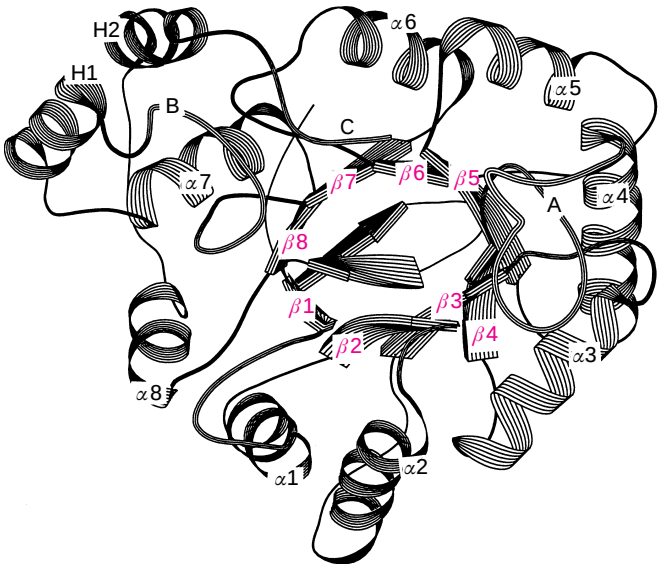


Figure 2 A common three-dimensional fold for the AKRs

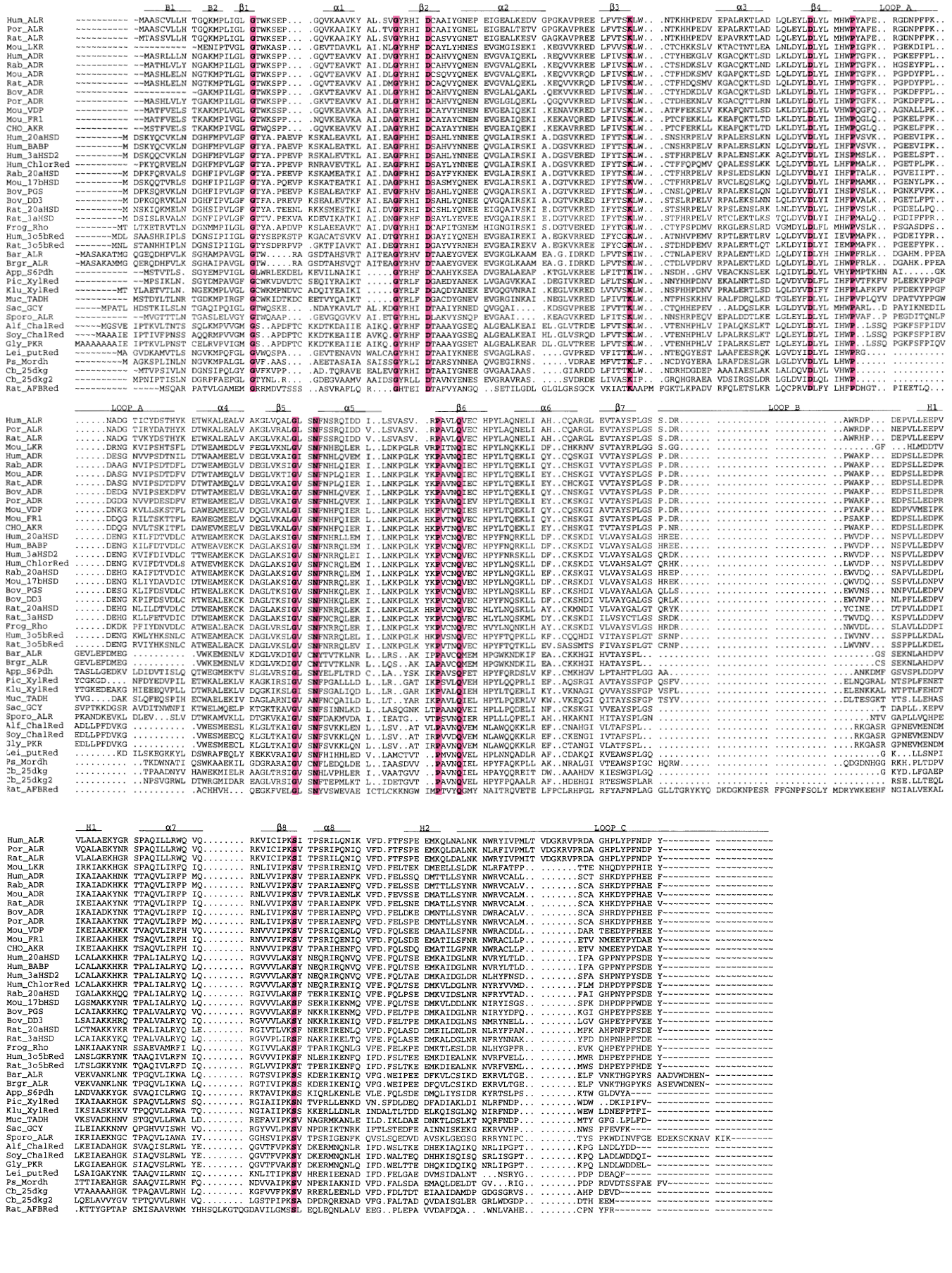
The ribbon drawing of the three-dimensional structure of rat liver 3α-HSD shows the common $(\alpha/\beta)_8$-barrel motif of the AKRs, with the C-terminus at the top. The α-helices and β-sheets (in pink) of the barrel are indicated, as are two helices (H1 and H2) not in the barrel. In addition, the A-loop, B-loop and C-terminal loop are indicated. Reproduced with permission from Hoog, Pawlowski, Alzari, Penning and Lewis (1994) Proc. Natl. Acad. Sci. U.S.A. **91**, 2517–2521 [9]; copyright (1994) National Academy of Sciences, U.S.A.

The structures of ADR, ALR, 3α-HSD and FR-1 indicate spatial conservation of the active-site residues Asp-50, Tyr-55, Lys-84 and His-117 (Figure 4) [7–10,69]. In addition, a water molecule is consistently bound between Tyr-55, His-117 and the C-4 of bound $NADP^+$, an arrangement thought to mimic the location of the substrate carbonyl. The spatial arrangement of Tyr-55, His-117 and the nicotinamide ring has been proposed to create an oxyanion hole. Based on the original ADR·NADPH binary complex structure, Wilson and co-workers [7] suggested that either Tyr-55 or His-117 is the proton donor, based on the proximity required between the C-4 position of the nicotinamide ring and the anticipated position of the substrate carbonyl. Moreover, Asp-50 and Lys-84 were proposed to form a hydrogen-bond/salt-bridge network that could decrease the pK_a of the tyrosine and facilitate proton donation. These authors also noted that the hydrophobic environment of His-117 would depress the pK_a of the imidazole side chain, making it less likely to function as the proton donor at physiological pH.

These structures laid the basis for site-directed-mutagenesis studies of the catalytic mechanism of ADR, ALR and 3α-HSD

Figure 3 Multiple sequence alignment of the AKR superfamily

The multiple sequence alignment used PILEUP from the GCG program suite [68] with the gap and the gap length penalties set to 12 and 4 respectively. The secondary structure of rat liver 3α-HSD is noted above the alignment [9]. B1, B2, H1 and H2 are β-sheets and α-helices not forming the core $(\alpha/\beta)_8$-barrel structure. Also, residues corresponding to three large loops on the C-terminal side of the barrel are noted. Invariant residues in all AKRs are highlighted in pink (see the text). Abbreviations for the proteins are given in Table 1. Modified from Jez et al. [26]; reproduced with the permission of Plenum Press.


β1 β2 β1 α1 β2 α2 β3 α3 β4 LOOP A
LOOP A α4 β5 α5 β6 α6 β7 LOOP B H1
H1 α7 β8 α8 H2 LOOP C
Hum_ALR
Por_ALR
Rat_ALR
Mou_LKR
Hum_ADR
Rab_ADR
Mou_ADR
Rat_ADR
Bov_ADR
Por_ADR
Mou_VDP
Mou_FR1
CHO_AKR
Hum_20aHSD
Hum_BABP
Hum_3aHSD2
Hum_ChlorRed
Rab_20aHSD
Mou_17bHSD
Bov_PGS
Bov_DD3
Rat_20aHSD
Rat_3aHSD
Frog_Rho
Hum_3o5bRed
Rat_3o5bRed
Bar_ALR
Brgr_ALR
App_S6Pdh
Pic_XylRed
Klu_XylRed
Muc_TADH
Sac_GCY
Sporo_ALR
Alf_ChalRed
Soy_ChalRed
Gly_PKR
Lei_putRed
Ps_Mordh
Cb_25dkg
Cb_25dkg2
Rat_AFBRed

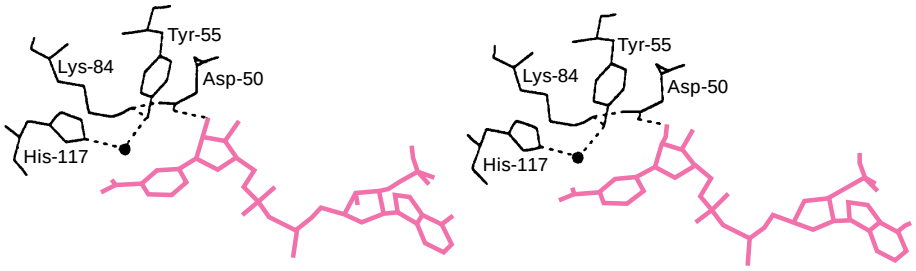


Figure 4 AKR active site

The three-dimensional structures of all AKRs retain the same spatial relationship between the active-site residues (Asp-50, Tyr-55, Lys-84 and His-117), the nicotinamide cofactor (shown in red) and a water molecule, as shown in this stereo-view of the rat liver 3α-HSD·NADP$^+$ binary-complex active site. Reprinted, with permission, from Bennett et al. [86]; copyright 1996 American Chemical Society.

[77–80]. In these enzymes, Asp-50 is not crucial for catalysis, although mutations of this residue interfere with substrate turnover and cofactor binding. Likewise, His-117 mutants of each enzyme retain some degree of catalytic activity, suggesting that this residue is also not the proton donor. Mutagenesis and molecular-modelling studies support a role for this histidine residue in orienting the substrate carbonyl in the active site [78,81]. Mutations of both Tyr-55 and Lys-84 severely impair or completely inactivate catalytic activity. However, based on the crystal structures, Lys-84 is not in the correct position to act as the general acid/base. The results to date thus favour Tyr-55 as the catalytic acid in the reaction mechanism [13].

Conservation of the active-site residues argues for a similar mechanism of oxidoreduction in other superfamily members. In all of the AKRs, both Asp-50 and Lys-84 are invariant. Tyr-55 is replaced only in ρ-crystallin, but this protein has extremely low enzymic activity, suggesting the importance of this residue in catalysis [82]. Interestingly, all the AKRs that catalyse carbonyl oxidoreduction reactions retain His-117, but the 3-oxo-5β-steroid 4-dehydrogenases do not.

The 3-oxo-5β-steroid 4-dehydrogenases catalyse the transfer of a hydride ion from the 4-*pro-R* position of NADPH to the 5β position on the steroid substrate, with the addition of an anti-facial proton at the 4α position [83]. This change in reaction mechanism is accompanied by a change of His-117 to a glutamic acid residue. As noted by Akhtar et al. [84], carbon–carbon double-bond reduction in steroids can occur by mechanisms that involve either protonation or hydride transfer occurring first. Protonation would give rise to a carbonium-ion intermediate, so that hydride addition could occur by a Markovnikov addition. The presence of a glutamate in the 3-oxo-5β-steroid 4-dehydrogenase active site could provide a mechanism for stabilizing an intermediate carbonium ion. If hydride transfer occurs first, then a negatively charged enolate would form at C-3 of the 3-oxo-5β-steroid substrate. Although this enolate could be stabilized by the oxyanion hole at the AKR active site, this hole may be compromised by the lack of His-117. Whether or not the equivalent of Tyr-55 would function as the general acid in double-bond reduction is unknown. Interestingly, an enolate is formed in 3-oxosteroid Δ^5–Δ^4-isomerase (Δ^5-3-ketosteroid isomerase) reaction, where a tyrosine does function as a general acid [85].

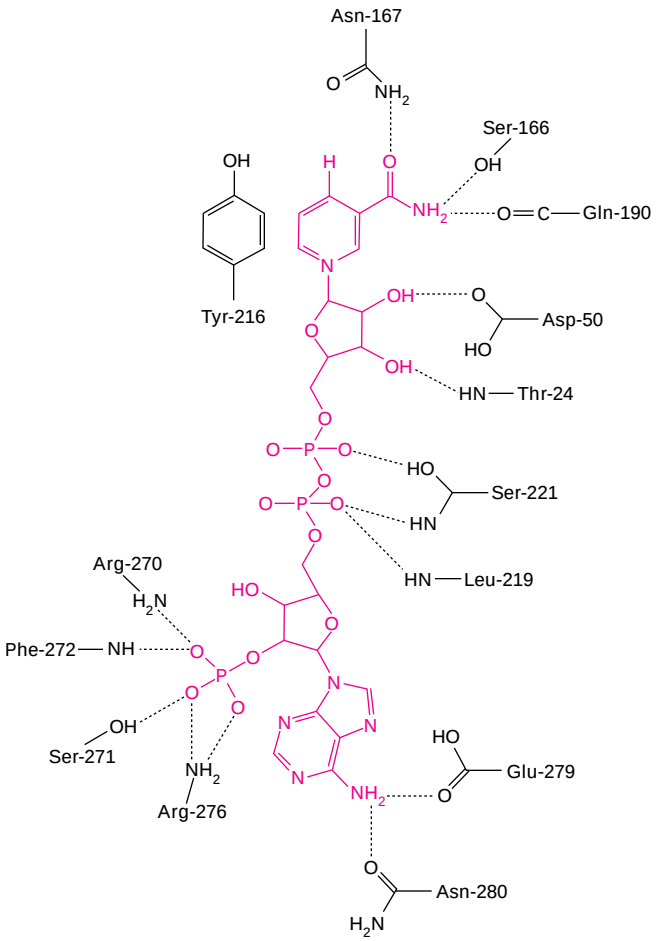


Figure 5 Schematic of NADP$^+$ binding in the AKRs

The atomic interactions between NADP$^+$ (depicted in red) and an AKR are indicated based on the structure of the 3α-HSD·NADP$^+$ binary complex [86]. Reprinted by permission of the publisher from Penning, Pawlowski, Schlegel, Jez, Lin, Smith-Hoog, Bennett and Lewis (1996) Steroids **61**, 508–523; copyright 1996 by Elsevier Science Inc.

NAD(P)(H) BINDING

The multiple sequence alignment, combined with the three-dimensional structures, suggests a similar mode of cofactor binding in all superfamily members. The enzyme·3NADP(H) binary-complex structures of ADR, ALR, FR-1 and 3α-HSD show the cofactor bound in an extended conformation on the C-terminal side of the barrel, with the nicotinamide ring located at the core of the barrel and the pyrophosphate bridge straddling the lip [7,8,10,69,86]. A short portion of the B-loop undergoes a

Table 2 Residues of the AKR cofactor-binding pocket

A multiple sequence alignment of the amino acids that comprise the cofactor binding pocket based on the known three-dimensional structures of AKRs is shown. Designations for the AKRs are as in Table 1. Strictly conserved residues are in bold face; positions with conservative changes are in italics.

Protein	Residue ...	24	50	166	167	190	216	219	221	270	271	272	276	279	280
Hum-ALR		W	**D**	*S*	**N**	**Q**	*Y*	*L*	S	*K*	**S**	I	*R*	*Q*	*N*
Por-ALR		W	**D**	*S*	**N**	**Q**	*Y*	*L*	S	*K*	**S**	V	*R*	*Q*	*N*
Rat-ALR		W	**D**	*S*	**N**	**Q**	*Y*	*L*	S	*K*	**S**	I	*R*	*Q*	*N*
Mou-LKR		W	**D**	*S*	**N**	**Q**	*Y*	*L*	G	*K*	**S**	V	*R*	*E*	*N*
Hum-ADR		W	**D**	*S*	**N**	**Q**	*Y*	*L*	S	*K*	**S**	V	*R*	*E*	*N*
Rab-ADR		W	**D**	*S*	**N**	**Q**	*Y*	*L*	S	*K*	**S**	V	*R*	*E*	*N*
Mou-ADR		W	**D**	*S*	**N**	**Q**	*Y*	*L*	S	*K*	**S**	V	*R*	*E*	*N*
Rat-ADR		W	**D**	*S*	**N**	**Q**	*Y*	*L*	S	*K*	**S**	V	*R*	*E*	*N*
Bov-ADR		W	**D**	*S*	**N**	**Q**	*Y*	*L*	S	*K*	**S**	V	*R*	*E*	*N*
Por-ADR		W	**D**	*S*	**N**	**Q**	*Y*	*L*	S	*K*	**S**	V	*R*	*E*	*N*
Mou-VDP		W	**D**	*S*	**N**	**Q**	*Y*	*L*	S	*K*	**S**	V	*R*	*E*	*N*
Mou-FR-1		W	**D**	*S*	**N**	**Q**	*Y*	*L*	S	*K*	**S**	V	*R*	*E*	*N*
CHO-AKR		W	**D**	*S*	**N**	**Q**	*Y*	*L*	S	*K*	**S**	V	*R*	*E*	*N*
Hum-20aHSD		Y	**D**	*S*	**N**	**Q**	*Y*	*L*	S	*K*	**S**	Y	*R*	*Q*	*N*
Hum-BABP		Y	**D**	*S*	**N**	**Q**	*Y*	*L*	S	*K*	**S**	Y	*R*	*Q*	*N*
Hum-3aHSDII		Y	**D**	*S*	**N**	**Q**	*Y*	*L*	S	*K*	**S**	Y	*R*	*Q*	*N*
Hum-ChlorRed		Y	**D**	*S*	**N**	**Q**	H	*L*	T	*K*	**S**	Y	*R*	*E*	*N*
Rab-20aHSD		Y	**D**	*S*	**N**	**Q**	*Y*	*L*	S	*K*	**S**	F	*R*	*E*	*N*
Mou-17bHSD		Y	**D**	*S*	**N**	**Q**	*Y*	*L*	S	*K*	**S**	F	*R*	*E*	*N*
Bov-Pgs		Y	**D**	*S*	**N**	**Q**	*Y*	*L*	A	*K*	**S**	F	*R*	*E*	*N*
Bov-DD3		F	**D**	*S*	**N**	**Q**	*Y*	*L*	S	*K*	**S**	Y	*R*	*E*	*N*
Rat-20aHSD		Y	**D**	*S*	**N**	**Q**	*Y*	*L*	T	*K*	**S**	F	*R*	*E*	*N*
Rat-3aHSD		T	**D**	*S*	**N**	**Q**	*Y*	*L*	S	*R*	**S**	F	*R*	*E*	L
Frog-Rho		Y	**D**	*S*	**N**	**Q**	*Y*	*L*	S	*K*	**S**	F	*R*	*Q*	*N*
Hum-3o5bRed		Y	**D**	*S*	**N**	**Q**	*Y*	*L*	T	*K*	**S**	F	*R*	*E*	*N*
Rat-3o5bRed		Y	**D**	*S*	**N**	**Q**	*Y*	*L*	T	*K*	**S**	T	*R*	*E*	*N*
Bar-ALR		W	**D**	*C*	**N**	**Q**	*Y*	*L*	–	*K*	**S**	S	*R*	*E*	*N*
Brgr-ALR		W	**D**	*C*	**N**	**Q**	*Y*	*L*	–	*K*	**S**	S	*R*	*E*	*N*
App-S6Pdh		W	**D**	*S*	**N**	**Q**	H	*L*	G	*K*	**S**	S	*R*	*E*	*N*
Pic-XylRed		W	**D**	*S*	**N**	**Q**	*Y*	F	P	*K*	**S**	N	*R*	*E*	*N*
Klu-XylRed		W	**D**	*S*	**N**	**Q**	*Y*	F	P	*K*	**S**	S	*R*	*D*	*N*
Muc-TADH		W	**D**	*A*	**N**	**Q**	*Y*	F	P	*K*	**S**	V	*R*	A	*N*
Sac-GCY		W	**D**	*S*	**N**	**Q**	*Y*	*L*	S	*K*	**S**	V	*R*	T	*N*
Sporo-ALR		W	**D**	*S*	**N**	**Q**	*Y*	*L*	N	*K*	**S**	V	*R*	*E*	*N*
Alf-ChalRed		–	**D**	*S*	**N**	**Q**	*F*	*L*	–	*K*	**S**	Y	*R*	*Q*	*N*
Soy-ChalRed		–	**D**	*S*	**N**	**Q**	*F*	*L*	–	*K*	**S**	Y	*R*	*Q*	*N*
Gly-PKR		–	**D**	*S*	**N**	**Q**	*F*	*L*	–	*K*	**S**	Y	*R*	*Q*	*N*
Lei-putRed		W	**D**	*S*	**N**	**Q**	*W*	*L*	Q	*K*	**S**	V	*R*	*E*	*N*
Ps-Morph		F	**D**	*C*	**N**	**Q**	*W*	*I*	C	*K*	**S**	V	*R*	K	*N*
Cb-25dkg		F	**D**	*S*	**N**	**Q**	*W*	*L*	Q	*K*	**S**	V	*R*	*E*	*N*
Cb2-25dkg		Y	**D**	*S*	**N**	**Q**	*W*	*L*	R	*K*	**S**	A	*R*	*E*	*N*
Rat-AFBRed		R	**D**	*S*	**N**	**Q**	*F*	*L*	G	S	**S**	L	*E*	L	A

conformational change upon cofactor binding and locks the pyrophosphate portion of NADP(H) in place; this represents the rate-limiting step in the reaction pathway [74,87,88].

The contacts between protein and cofactor are nearly identical in all known AKR structures (Figure 5). The nicotinamide ring and Tyr-216 stack so that the cofactor faces the active site to maintain 4-*pro-R* hydrogen transfer. In this orientation, the nicotinamide ring is in the *anti* conformation with respect to the ribose. This conformation is also dependent upon the hydrogen bonds made by Ser-166, Asn-167 and Gln-190 with the carboxamide moiety of the cofactor. Asp-50 and Thr-24 interact with the nicotinamide ribose, while the pyrophosphate backbone hydrogen-bonds with Leu-219 and Ser-221 of the B-loop. Finally, the adenosine 2′-monophosphate has extensive hydrogen-bond and charge–charge interactions with the protein through Arg-270, Ser-271, Phe-272, Arg-276, Glu-279 and (generally) Asn-280.

Mutagenesis studies of ADR and ALR demonstrate the importance of Arg-270 and Arg-276 in discriminating between the binding of NAD(H) and NADP(H) [89–91]. Additional mutagenesis has probed the contributions of other residues in cofactor binding. In ADR, a Y216F mutant had only slightly decreased affinity for the cofactor, consistent with the conservation of planar residues at this position that can stack with the nicotinamide ring [92]. Mutations of residues in contact with the nicotinamide ribose of the bound cofactor (Asp-50 and Thr-24) alter NADPH binding modestly [77,80,92]. However, Lys-84 mutants also exhibit changes in cofactor affinities, possibly reflecting alterations in the hydrogen-bond network with Asp-50 [77,80].

The identity of the cofactor-binding pocket across all the AKRs is striking, being evident even between proteins of less than 30% overall amino acid sequence identity (Table 2). One-third of the residues of the cofactor pocket are strictly conserved (Asp-50, Asn-167, Gln-190 and Ser-271), and more are highly conserved (Ser-166, Leu-219, Arg-270, Arg-276, Glu-279 and Asn-280), implying a similar mode of cofactor binding across the superfamily. Also, a planar amino acid is always found at

position 216, and would maintain the stacking interaction with the nicotinamide ring. Since the residues involved in NAD(P)(H) binding encompass regions from the centre to the periphery of the $(\alpha/\beta)_8$-barrel structure, the conservation of the overall three-dimensional fold and of the cofactor-binding site may be related.

SUBSTRATE SPECIFICITY

Based on the locations of the active site and the cofactor-binding pocket in known AKR structures, a putative substrate-binding site has been proposed [9,86]. The topology of the substrate binding site resembles a cleft, and includes three components (Figure 6): (1) an oxyanion binding site (Tyr-55, His-117 and C-4 of the nicotinamide ring); (2) residues at the edge of the active site (Ala-52, Leu-54, Trp-86 and Phe-118); and (3) amino acids from three loops forming the sides of the cleft. The A-loop (Met-120, Phe-128 and Phe-129) contributes to one side of the apolar cleft, with the opposite side being formed by the B-loop (Trp-227) and the C-terminal loop (Asn-306, Ala-308 and Tyr-310).

Functional studies have confirmed the basic structural components of the substrate binding site in the AKRs. Of the residues near the active site, mutagenesis of Trp-86 in 3α-HSD modestly affects steroid binding; in the B-loop, mutation of Trp-227 severely impairs steroid binding [21]. Likewise, in the C-loop, point and deletion mutants of ADR and ALR hinder substrate binding, indicating the importance of this loop in forming part of the apolar cleft [80,93–95]. However, structure/function studies are only now addressing the following two questions: (1) how do the AKRs distinguish between sugar and steroid substrates? and (2) how do the HSDs attain their exquisite positional and stereo-specificity for individual steroids?

A comparison of the residues in the putative substrate-binding pocket shows that some portions are variable, while others are conserved (Table 3). In general, positions near the catalytic residues tend to be conserved, with greater variation occurring in the loops. The active-site tyrosine and histidine are conserved for catalytic purposes, but could also lock the carbonyl group of the substrate into position for hydride transfer. Functional analysis supports the importance of these amino acids in substrate binding, since removal of the Tyr-55 hydroxy group (a Y55F mutant) produces a 30-fold increase in the K_d for testosterone in 3α-HSD [79]. Likewise, His-117 mutants of ADR result in an increase in the K_m for D-glyceraldehyde of almost 80000-fold [77] and can no longer discriminate between the D- and L-isomers of xylose [78].

Given the structures of aldehyde, monosaccharide and steroid substrates, it becomes apparent that residues at the edge of the active site may discriminate between these molecules. Near the active site, Ala-52 and Trp-86 are conserved. However, amino acids at positions 54 and 118 vary depending upon the substrate, and may have a role in determining sugar versus steroid specificity. Residue 54 is generally leucine or isoleucine in the HSDs, and is valine in the ADRs. Differences in this residue may alter the binding-site topology to accommodate different sized substrates. Also, in the ALRs and ADRs residue 118 is tryptophan, whereas it is nearly always phenylalanine in the HSDs, and is either valine or methionine in the 5β-reductases. Molecular dynamics studies of ADR suggest that the indole nitrogen of Trp-118 hydrogen-bonds with the 2′-hydroxy group of substrates containing an R-CHOH-CHO moiety [81]. Also, ADR·NADPH·inhibitor ternary-complex structures exhibit this interaction [96,97]. Replacement of the tryptophan by phenylalanine or valine would remove this hydrogen bond and possibly provide a degree of discrimination between sugar and steroid substrates by controlling accessibility to or substrate orientation within the active site.

Beyond the immediate vicinity of the active site, differences in the composition and size of the three loops on the C-terminal side of the barrel suggest their role as additional determinants of substrate specificity. Since steroids are much larger than monosaccharides, these loops could determine the positional and stereo-specificity observed in the HSDs of the superfamily.

We have carried out homology modelling of the structures of mouse liver 17β-HSD and rat ovary 20α-HSD (EC 1.1.1.149) using the three-dimensional co-ordinates of rat liver 3α-HSD as a template. The resulting models demonstrate the conservation of the barrel structure and highlight sequence differences within the three loops on the C-terminal side of the barrel (Figure 7). Docking of the steroid in the apolar cleft of rat liver 3α-HSD, based on mutagenesis results, suggests that the C-3 carbonyl of the steroid substrate is bound at the catalytic tetrad, with the α-face of the steroid oriented towards the A-loop, preserving the stereochemistry of hydride transfer [21]. In the case of 17β-HSD, the steroid would have to bind backwards to place the substrate

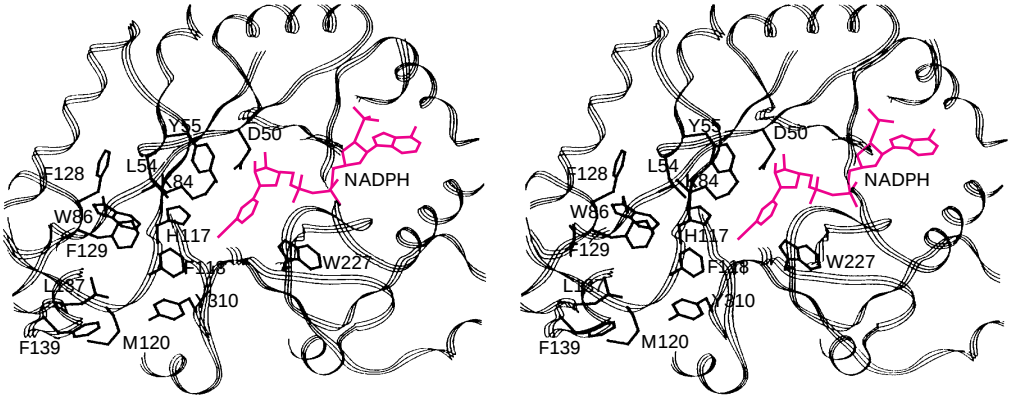


Figure 6 Substrate binding site

This stereo-view displays the residues of the proposed substrate binding site in relation to the active-site residues (Asp-50, Tyr-55, Lys-84 and His-117) and bound $NADP^+$ (depicted in red) of 3α-HSD [86]. Apolar residues near the active site (Leu-54, Trp-86 and Phe-118) and on the A-loop (including Met-120, Phe-128, Phe-129, Leu-137 and Phe-139) are shown. The location of Trp-227 on the B-loop, and the position of Tyr-310 on the C-terminal loop are also indicated. Figure prepared using SETOR [100].

Table 3 Residues of the AKR substrate pocket

A multiple sequence alignment of the amino acids that comprise the putative substrate binding pocket based on the known three-dimensional structures for AKRs is shown. Designations of the AKRs are as in Table 1. Highly conserved residues with a proposed role in substrate binding are indicated in bold face.

Protein	Residue ...	52	54	55	86	117	118	120	128	129	227	306	308	310
Hum-ALR		**A**	I	**Y**	**W**	**H**	W	Y	P	F	W	I	P	R
Por-ALR		**A**	I	**Y**	**W**	**H**	W	Y	P	F	W	I	P	R
Rat-ALR		**A**	V	**Y**	**W**	**H**	W	Y	P	F	W	I	P	R
Mou-LKR		**A**	L	**Y**	**W**	**H**	W	I	D	I	–	A	F	T
Hum-ADR		**A**	V	**Y**	**W**	**H**	W	T	F	F	W	C	L	S
Rab-ADR		**A**	V	**Y**	**W**	**H**	W	T	Y	F	W	C	L	S
Mou-ADR		**S**	V	**Y**	**W**	**H**	W	T	Y	F	W	C	L	S
Rat-ADR		**A**	V	**Y**	**W**	**H**	W	T	Y	F	W	C	L	S
Bov-ADR		**A**	V	**Y**	**W**	**H**	W	T	F	F	W	C	L	S
Por-ADR		**A**	V	**Y**	**W**	**H**	W	T	P	F	W	C	L	S
Mou-VDP		**A**	V	**Y**	**W**	**H**	W	Q	L	L	Y	C	L	D
Mou-FR-1		**A**	A	**Y**	**W**	**H**	W	Q	L	F	S	C	L	E
CHO-AKR		**A**	A	**Y**	**W**	**H**	W	Q	L	F	W	C	L	E
Hum -20aHSD		**A**	L	**Y**	**W**	**H**	F	V	V	I	W	L	L	I
Hum-BABP		**A**	V	**Y**	**W**	**H**	F	V	V	I	W	L	L	I
Hum-3aHSDII		**A**	L	**Y**	**W**	**H**	S	M	L	S	W	F	S	S
Hum-ChlorRed		**A**	L	**Y**	**W**	**H**	F	M	P	L	W	V	M	F
Rab-20aHSD		**A**	F	**Y**	**W**	**H**	F	T	I	I	W	V	A	F
Mou-17bHSD		**A**	M	**Y**	**W**	**H**	F	M	Y	L	W	I	G	S
Bov-Pgs		**A**	L	**Y**	**W**	**H**	S	V	F	V	W	Y	F	K
Bov-DD3		**A**	L	**Y**	**W**	**H**	F	V	L	F	W	N	L	L
Rat-20aHSD		**S**	L	**Y**	**W**	**H**	F	V	L	L	C	F	A	M
Rat-3aHSD		**A**	L	**Y**	**W**	**H**	F	M	F	F	W	N	A	Y
Frog-Rho		**A**	I	**T**	**W**	**H**	W	V	S	D	W	G	F	E
Hum-3o5bRed		**A**	I	**Y**	**W**	**E**	V	M	I	Y	W	V	L	M
Rat-3o5bRed		**A**	V	**Y**	**W**	**E**	M	M	F	Y	W	V	M	M
Bar-ALR		**A**	E	**Y**	**W**	**H**	W	F	M	–	–	L	G	E
Brgr-ALR		**A**	E	**Y**	**W**	**H**	W	F	K	–	–	L	G	E
App-S6Pdh		**A**	H	**Y**	**W**	**H**	Y	M	–	–	–	S	P	K
Pic-XylRed		**A**	D	**Y**	**W**	**H**	F	V	K	Y	L	N	P	W
Klu-XylRed		**A**	D	**Y**	**W**	**H**	F	L	K	Y	L	N	P	W
Muc-TADH		**A**	D	**Y**	**W**	**H**	F	V	V	Y	L	N	P	M
Sac-GCY		**A**	I	**Y**	**W**	**H**	W	A	K	N	–	V	P	N
Sporo-ALR		**A**	V	**Y**	**W**	**H**	W	V	I	T	–	N	P	T
Alf-ChalRed		**A**	A	**Y**	**W**	**H**	W	–	S	F	–	P	P	K
Soy-ChalRed		**A**	A	**Y**	**W**	**H**	W	–	S	F	–	S	P	K
Gly-PKR		**A**	A	**Y**	**W**	**H**	W	–	S	F	–	P	P	K
Lei-putRed		**A**	I	**Y**	**W**	**H**	W	R	–	–	–	S	Y	P
Ps-Morph		**A**	S	**Y**	F	**H**	W	–	–	–	D	–	I	P
Cb-25dkg		**A**	I	**Y**	**W**	**H**	W	–	–	–	–	G	V	A
Cb2-25dkg		**A**	N	**Y**	P	**H**	W	–	–	–	–	D	D	D
Rat-AFBRed		**A**	V	**Y**	A	**H**	F	D	T	L	Y	V	H	C

carbonyl at C-17 towards the catalytic residues, and upside-down with the β-face towards the A-loop to maintain the observed stereochemistry. For 20α-HSD, the steroid substrate need only bind backwards. These different binding modes may be permitted by alterations in the A-loop. In 3α-HSD, Phe-128 and Phe-129 are in the A-loop. By contrast, in 17β-HSD, these residues are replaced by tyrosine and leucine respectively, and 20α-HSD has two leucines in these positions. These differences would retain the apolar nature of the proposed binding site, but may generate steroid-specific van der Waals interactions that contribute to the positional and stereo-specificity of the reaction. In addition, the B- and C-loops on the opposite side of the apolar cleft could provide further specificity in the HSDs of the superfamily.

The atomic interactions that exist between the substrate and the B- and C-loops are difficult to model, because these loops are the least well defined features of the three-dimensional structure of rat liver 3α-HSD [9,86]. Whereas mutation of Trp-227 on the B-loop dramatically decreases the affinity for testosterone binding [21], this residue is highly conserved across the superfamily, implying that it does not play a role in substrate selectivity. It is possible that other amino acids on the B-loop interact with the substrate. Finally, the role of the C-loop in steroid binding remains unclear.

Residues of the C-terminal loop are found near the active site (position 306) and further along the apolar cleft (positions 308 and 310), suggesting a role for this loop in discriminating between sugars and steroids and also in defining steroid specificity. Among the HSDs this region displays a large degree of sequence variation, with no clear pattern relating to substrate specificity. However, in the ADRs residues 306, 308 and 310 are highly conserved, with a Cys-Leu-Ser sequence; in all the mammalian ALRs these residues become Ile-Pro-Leu/Ile, and in the plant ALRs they are Leu-Gly-Glu. Mutations of position 306 in ADR and ALR affect catalytic efficiency with a number of substrates [80,93]. Also, deletion of the C-terminal loop in ADR after position 295 selectively affects catalytic efficiency with uncharged substrates (e.g. D,L-glyceraldehyde and D-glucose),

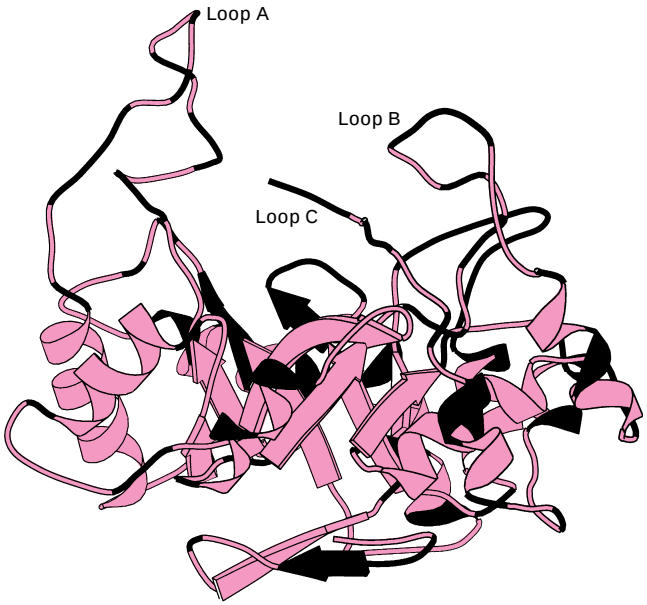


Figure 7 Homology modelling of 17β- and 20α-HSDs using 3α-HSD as a template: the three C-terminal loops

The crystallographic co-ordinates of the rat liver 3α-HSD apoenzyme structure (PDB entry 1RAL; [9]) were used to perform homology modelling of mouse liver 17β-HSD [27] and rat ovary 20α-HSD [28] using QUANTA (Molecular Simulations Inc., Waltham, MA, U.S.A.). Since both 17β- and 20α-HSD are of similar size and share high ($>$ 70%) amino acid sequence identity with 3α-HSD, the polypeptide backbone and homologous residues were modelled directly from 3α-HSD, with dissimilar residues built into the model structures using standard Karplus rotomers. The initial models were energy-minimized using a conjugate gradient algorithm for 200 cycles with CHARMm (Molecular Simulations Inc.). The resulting structures were overlaid to determine the spatial relationship between the three proteins. Pink regions of the structure correspond to segments of high sequence and structural identity, and black regions correspond to areas of low sequence and structural identity. This view shows the $(\alpha/\beta)_8$-barrel from the side and is rotated 180 ° about the barrel axis relative to the view shown in Figure 2. Figure prepared with MolScript [101].

but has little effect on the k_{cat}/K_m value for charged substrates (e.g. D-glucuronate) [94]. Recently, Barski et al. [95] have shown that the arginine corresponding to position 310 in the C-terminal loop of ALR determines the specificity of ALR for substrates with a negatively charged carboxylate group, such as glucuronate, succinic semialdehyde and *p*-carboxybenzaldehyde. Together, these results indicate that the C-terminal loop can play an important role in determining substrate specificity in the proteins of the AKR superfamily.

EVOLUTIONARY RELATIONSHIPS IN THE AKR SUPERFAMILY

Since members of the superfamily show marked sequence identity and occur in species from protozoa to mammals, the AKR superfamily may represent a case of divergent evolution from a multifunctional ancestor protein. Too little is known about the gene structures of the AKRs to assess the mechanism of divergence, but the common constellation of active-site residues and the highly conserved NAD(P)(H)-binding pocket implies gene duplication and subsequent evolution of substrate specificity.

The substrate specificity of these proteins is also consistent with divergence from a common ancestor. Although common substrates, such as 4-nitrobenzaldehyde, are turned over by the AKRs, each AKR can also display unique substrate specificity. The 3α-HSD/DD isoforms of rat and human liver provide an example of such divergence. As the only form found in rat liver, the rat enzyme acts as a 3α-HSD/DD and a bile acid-binding protein [11]. This enzyme is 65% identical to the human liver 3α-HSD/DD isoforms (DD2 and DD4), but each human isoform has unique functions. DD2 and DD4 both have 3α-HSD/DD activity and are 86% identical with each other in amino acid sequence. However, DD4 has a high 3α-HSD activity and does not bind bile acids, while DD2 has low 3α-HSD activity and binds bile acids with high affinity [31]. Similar arguments for gene duplication and divergence of substrate specificity have been proposed for the HSDs of the SDR superfamily [22].

Based on the results of functional studies and our multiple sequence alignment, we suggest that divergence of function in the AKRs occurs through modification of the three large loops on the C-terminal side of the barrel, with conservation of the $(\alpha/\beta)_8$-barrel architecture to maintain the active site and the NAD(P)(H)- binding pocket. A precedent for this exists in glucose oxidase and cholesterol oxidase, where the evolution of similar enzymes to catalyse the same reaction on either sugars or steroids has been observed [98]. Reminiscent of the AKRs, these proteins have the same three-dimensional fold and similar cofactor-binding domains, as well as differences in the loops that alter substrate accessibility to the active site.

In addition to the divergent evolution of substrate specificity, the AKR superfamily may also be an example of convergent evolution to the same constellation of active-site residues as the SDRs. Although the AKRs and SDRs are not homologous and have completely different three-dimensional folds, the catalytic mechanisms of the two superfamilies are analogous. The AKRs catalyse NAD(P)(H)-dependent oxidoreduction reactions, in which a tyrosine and a lysine are key catalytic components. Likewise, the SDRs are nicotinamide-dependent enzymes defined by a Tyr-Xaa-Xaa-Xaa-Lys motif at the active site [2]. An overlay of the active sites of an SDR (3α/20β-HSD from *Streptomyces hydrogenans*) and an AKR (rat liver 3α-HSD) showed that the tyrosine hydroxy group of each protein differs in space by less than 0.5 Å, while the lysines maintain approximately the same location relative to the hydroxy group of the tyrosine [86]. As in the classic example of the serine proteases, the evolution of similar active sites may be a consequence of a single energetically accessible pathway for a given chemical mechanism [99]. Although mechanistic studies of the AKRs and the SDRs have not progressed to the same level as studies of the serine proteases, these initial results are intriguing.

CONCLUSION

The AKR superfamily represents an interesting challenge to understanding structure/function relationships, for two main reasons. First, the mammalian AKRs represent potential drug targets, but the similarity of these proteins has made the design of specific inhibitors difficult. Only by understanding the structure/function relationship of the superfamily will rational drug design against individual AKRs be possible. Although the catalytic mechanism and mode of NAD(P)(H) binding in the AKRs is understood, how these proteins discriminate between sugars and steroids, and how positional and stereo-selectivity between different steroids is achieved, remain to be elucidated. Secondly, the architecture of these proteins may prove useful for studying the evolution of substrate specificity and of how a common scaffold can catalyse a similar reaction with a broad range of substrates. In the future, it may be possible to use the AKR scaffold as a basis for engineering new substrate specificity by altering the C-terminal loops of the barrel.

This work was supported by NIH grants DK47015 and CA39504 (to T.M.P.), and by a fellowship from the Cancer Research Fund of the Damon Runyon–Walter Winchell Foundation, DRG-1298 (to M.J.B. and M.L.). A preliminary account of this work was presented at the 2nd International Symposium on Molecular Steroidogenesis in Monterey, CA, U.S.A., between 8 and 11 June, 1996.

REFERENCES

1 Jornvall, H., Persson, M. and Jeffrey, J. (1981) Proc. Natl. Acad. Sci. U.S.A. **78**, 4226–4230
2 Jornvall, H., Persson, M., Krook, M., Atrian, S., Gonzalez-Duarte, R., Jeffrey, J. and Ghosh, D. (1995) Biochemistry **34**, 6003–6013
3 Eklund, H., Horjales, E., Jornvall, H., Branden, C. I. and Jeffrey, J. (1985) Biochemistry **24**, 8005–8012
4 Ghosh, D., Weeks, C. M., Grochulski, P., Duax, W. L., Erman, M., Rimsay, R. L. and Orr, J. C. (1991) Proc. Natl. Acad. Sci. U.S.A. **88**, 10064–10068
5 Bohren, K. M., Bullock, B., Wermuth, B. and Gabbay, K. H. (1989) J. Biol. Chem. **264**, 9547–9551
6 Bruce, N. C., Willey, D. L., Coulson, A. F. W. and Jeffrey, J. (1994) Biochem. J. **299**, 805–811
7 Wilson, D. K., Bohren, K. M., Gabbay, K. H. and Quiocho, F. A. (1992) Science **257**, 458–460
8 Wilson, D. K., Nakano, T., Petrash, J. M. and Quiocho, F. A. (1995) Biochemistry **34**, 14323–14330
9 Hoog, S. S., Pawlowski, J. E., Alzari, P. M., Penning, T. M. and Lewis, M. (1994) Proc. Natl. Acad. Sci. U.S.A. **91**, 2517–2521
10 El-Kabbani, O., Judge, K., Ginell, S. L., Myles, D. A. A., DeLucas, L. J. and Flynn, T. G. (1995) Nature Struct. Biol. **2**, 687–692
11 Penning, T. M., Pawlowski, J. E., Schlegel, B. P., Jez, J. M., Lin, H.-K., Smith-Hoog, S., Bennett, M. J. and Lewis, M. (1996) Steroids **61**, 508–523
12 Welle, R., Schroder, G., Schiltz, E., Grisebach, H. and Schroder, J. (1991) Eur. J. Biochem. **196**, 423–430
13 Petrash, J. M., Tarle, I., Wilson, D. K. and Quiocho, F. A. (1995) Diabetes **43**, 955–959
14 Tomlinson, D. R., Stevens, E. J. and Diemel, L. T. (1994) Trends Pharmacol. Sci. **15**, 293–297
15 Kador, P. F. and Sharpless, N. E. (1983) Mol. Pharmacol. **24**, 521–531
16 Turner, A. J., Illingworth, J. A. and Tipton, K. F. (1974) Biochem. J. **144**, 353–360
17 Penning, T. M. and Talalay, P. (1983) Proc. Natl. Acad. Sci. U.S.A. **80**, 4504–4508
18 Watanabe, K., Fujii, Y., Nakayama, K., Ohkubo, H., Kuramitsu, S., Kagamiyama, H., Nakanishi, S. and Hayaishi, O. (1988) Proc. Natl. Acad. Sci. U.S.A. **85**, 11–15
19 Dorfman, R. I. and Shipley, R. A. (1956) Androgens, pp. 122–126, John Wiley & Sons, New York
20 Lacy, W. R., Washenick, K. J., Cook, R. G. and Dunbar, B. S. (1993) Mol. Endocrinol. **7**, 58–66
21 Jez, J. M., Schlegel, B. P. and Penning, T. M. (1996) J. Biol. Chem. **271**, 30190–30198
22 Baker, M. E. (1996) Bioessays **18**, 63–70
23 Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990) J. Mol. Biol. **215**, 403–410
24 Gish, W. and States, D. J. (1993) Nature Genet. **3**, 266–272
25 Pawlowski, J. E., Huizinga, M. and Penning, T. M. (1991) J. Biol. Chem. **266**, 8820–8825
26 Jez, J. M., Flynn, T. G. and Penning, T. M. (1996) in Enzymology and Molecular Biology of Carbonyl Metabolism, vol. 6, (Weiner, H., Lindahl, R., Crabb, D. W. and Flynn, T. G., eds.), pp. 579–589, Plenum Press, New York
27 Deyashiki, Y., Ohshima, K., Nakanishi, M., Sato, K., Matsuura, K. and Hara, A. (1995) J. Biol. Chem. **270**, 10461–10467
28 Miura, R., Shiota, K., Noda, K., Yagi, S., Ogawa, T. and Takahashi, M. (1994) Biochem. J. **299**, 561–567
29 Hara, A., Matsuura, K., Tamada, Y., Sato, K., Miyabe, Y., Deyashiki, Y. and Ishida, N. (1996) Biochem. J. **313**, 373–376
30 Stolz, A., Hammond, L., Lou, H., Takikawa, H., Ronk, M. and Shively, J. E. (1993) J. Biol. Chem. **268**, 10448–10457
31 Deyashiki, Y., Ogasawara, A., Nakayama, T., Nakanishi, M., Miyabe, Y., Sato, K. and Hara, A. (1994) Biochem. J. **299**, 545–552
32 Khanna, M., Qin, K. N., Wang, R. W. and Cheng, K. C. (1995) J. Biol. Chem. **270**, 20162–20168
33 Winters, C. J., Molowa, D. T. and Guzelian, P. S. (1990) Biochemistry **29**, 1080–1087
34 Chung, S. and LaMedola, J. (1989) J. Biol. Chem. **264**, 14775–14777
35 Garcia-Perez, A., Martin, B., Murphy, H. R., Uchida, S., Murer, H., Cowley, B. D., Handler, J. S. and Burg, M. B. (1989) J. Biol. Chem. **264**, 16815–16821
36 Gui, T., Tanimoto, T., Kokai, Y. and Nishimura, C. (1995) Eur. J. Biochem. **227**, 448–453
37 Carper, D., Nishimura, C., Shinohara, T., Dietzchold, B., Wistow, G., Craft, C., Kador, P. and Kinoshita, J. H. (1987) FEBS Lett. **220**, 209–213
38 Petrash, J. M. and Favello, A. D. (1989) Exp. Eye Res. **8**, 1021–1027
39 Warren, J. C., Murdock, G. L., Ma, Y., Goodman, S. R. and Zimmer, W. E. (1993) Biochemistry **32**, 1401–1406
40 Jaquinod, M., Potier, R., Klarskov, K., Reymann, J. M., Sorokine, O., Kieffer, S., Barth, P., Andriantomanga, V., Biellman, J. F. and Van Dorsselaes, A. (1993) Eur. J. Biochem. **218**, 893–903
41 Bohren, K. M., Barski, O. A. and Gabbay, K. H. (1996) in Enzymology and Molecular Biology of Carbonyl Metabolism, vol. 6, (Weiner, H., Lindahl, R., Crabb, D. W. and Flynn, T. G., eds.), (Adv. Exp. Med. Biol. **414**), pp. 455–464, Plenum Press, New York
42 Wermuth, B., Omar, A., Forster, A., di Francesco, G., Wolf, M., von Wartbuth, J. P., Bullock, B. and Gabbay, K. H. (1987) Prog. Clin. Biol. Res. **232**, 297–307
43 Flynn, T. G., Green, N. C., Bhatia, M. B. and El-Kabbani, O. (1995) Adv. Exp. Med. Biol. **372**, 193–201
44 Takahasi, M., Fujii, J., Teshima, T., Suzuki, K., Shiba, T. and Taniguchi, N. (1993) Gene **127**, 249–253
45 Bartels, D., Engelhardt, K., Roncarati, R., Schneider, K., Rotter, M. and Salamini, F. (1991) EMBO J. **10**, 1037–1043
46 Lee, S. P. and Chen, T. H. H. (1993) Plant Physiol. **101**, 1089–1092
47 Kita, K., Matsuzaki, K., Hashimoto, T., Yanase, H., Kato, N., Chung, M. C. M., Kataoka, M. and Shimizu, S. (1996) Appl. Environ. Microbiol. **62**, 2303–2310
48 Kondo, K., Kai, M., Setoguchi, Y., Eggertsen, G., Sjoblom, P., Setoguchi, T., Okuda, K. and Bjorkhem, I. (1994) Eur. J. Biochem. **219**, 357–363
49 Onishi, Y., Noshiro, M., Shimosato, T. and Okuda, K. (1994) FEBS Lett. **283**, 215–218
50 Sallaud, C., El Turk, C., Bigerre, L., Sevin, H., Welle, R. and Esnault, R. (1995) Plant Physiol. **108**, 869–870
51 Akashi, T., Furuno, T., Futami, K., Honda, M., Takahashi, T., Welle, R. and Ayabe, S. (1996) Plant Physiol. **111**, 347
52 Anderson, S., Berman-Marks, C., Lazarus, R., Miller, J., Stafford, K., Seymour, J., Light, D., Rastetter, W. and Estell, D. (1985) Science **230**, 144–149
53 Grindley, J. G., Payton, M. A., De Pol, H. and Hardy, K. G. (1988) Appl. Environ. Microbiol. **54**, 1770–1775
54 Willey, D. L., Caswell, D. A., Lowe, C. R. and Bruce, N. C. (1993) Biochem. J. **290**, 539–544
55 Samaras, N. and Spithill, T. W. (1989) J. Biol. Chem. **264**, 4251–4254
56 Terada, T., Adachi, H., Nanjo, H., Fujita, N., Takagi, T., Nishikawa, J., Imagawa, M., Nishihara, T. and Maeda, M. (1996) in Enzymology and Molecular Biology of Carbonyl Metabolism, vol. 6, (Weiner, H., Lindahl, R., Crabb, D. W. and Flynn, T. G., eds.), (Adv. Exp. Med. Biol. **414**), pp. 545–553, Plenum Press, New York
57 Tomarev, S. I., Zinovieva, R. D., Dolgilevich, S. M., Luchin, S. V., Krayev, A. S., Skyryabin, K. G. and Gause, G. G. (1984) FEBS Lett. **171**, 297–301
58 Pailhoux, E. A., Martinez, A., Veyssiere, G. M. and Jean, C. G. (1990) J. Biol. Chem. **265**, 19932–19936
59 Kanayama, Y., Moni, H., Imaseki, H. and Yamaki, S. (1992) Plant Physiol. **100**, 1607–1608
60 Amore, R., Kotter, P., Kuster, C., Ciriacy, M. and Hollenberg, C. P. (1991) Gene **109**, 89–97
61 Billard, P., Menart, S., Fleer, R. and Bolotin-Fukuhara, M. (1995) Gene **162**, 93–97
62 Czempinski, K., Kruft, V., Wostermeyer, J. and Burmester, A. (1996) Microbiology **142**, 2647–2654
63 Oechsner, U., Magodolen, V. and Bandlow, W. (1988) FEBS Lett. **238**, 123–128
64 Hyndman, D. J., Takenoshita, R., Vera, N. L., Pang, S. G. and Flynn, T. G. (1997) J. Biol. Chem. **272**, 13286–13291
65 Donohue, P. J., Alberts, G. F., Hampton, B. S. and Winkles, J. A. (1994) J. Biol. Chem. **269**, 8604–8609
66 Ellis, E. M., Judah, D. J., Neal, G. E. and Hayes, J. D. (1993) Proc. Natl. Acad. Sci. U.S.A. **90**, 10350–10354
67 McCormack, T. and McCormack, K. (1994) Cell **79**, 1133–1135
68 Devereux, J., Haeberil, P. and Smithies, O. (1984) Nucleic Acids Res. **12**, 387–395
69 Rondeau, J. M., Tete-Favier, F., Podjarny, A., Reymann, J. M., Barth, P., Biellman, J. F. and Moras, D. (1992) Nature (London) **355**, 469–472
70 Scheerlink, J. P., Lasters, I., Claessens, M., De Maeyer, M., Pio, F., Delhaise, P. and Wodak, S. J. (1992) Proteins Struct. Funct. Genet. **12**, 299–313
71 Janecek, S. (1996) Protein Sci. **5**, 1136–1142
72 Farber, G. K. and Petsko, G. A. (1990) Trends Biochem. Sci. **15**, 228–234
73 Davidson, W. S. and Flynn, T. G. (1979) Biochem. J. **177**, 595–601
74 Grimshaw, C. E., Shahbaz, M. and Putney, C. G. (1990) Biochemistry **29**, 9947–9955
75 Askonas, L. J., Ricigliano, J. W. and Penning, T. M. (1991) Biochem. J. **278**, 835–841
76 Yoshida, M., Kuroki, Y., Kobayashi, E. and Tamaoki, B. (1992) J. Steroid Biochem. Mol. Biol. **41**, 29–36

77 Tarle, I., Borhani, D. W., Wilson, D. K., Quiocho, F. A. and Petrash, J. M. (1993) J. Biol. Chem. **268**, 25687–25693
78 Bohren, K. M., Grimshaw, C. E., Lai, C. J., Harrison, D. H., Ringe, D., Petsko, G. A. and Gabbay, K. H. (1994) Biochemistry **33**, 2021–2032
79 Pawlowski, J. E. and Penning, T. M. (1994) J. Biol. Chem. **269**, 13502–13510
80 Barski, O. A., Gabbay, K. H., Grimshaw, C. E. and Bohren, K. M. (1995) Biochemistry **34**, 11264–11272
81 De Winter, H. L. and von Itzstein, M. (1995) Biochemistry **34**, 8299–8308
82 Fujii, Y., Watanabe, K., Hayashi, H., Urade, Y., Kuramitsu, S., Kagamiyama, H. and Hayaishi, O. (1990) J. Biol. Chem. **265**, 9914–9923
83 Berseus, O. and Bjorkhem, I. (1967) Eur. J. Biochem. **2**, 503–507
84 Akhtar, M., Wilton, D. C., Watkinson, I. A. and Rahimtula, A. D. (1972) Proc. R. Soc. London B **180**, 157–177
85 Zhao, Q., Abeygunawardana, C., Talalay, P. and Mildvan, A. S. (1996) Proc. Natl. Acad. Sci. U.S.A. **93**, 8220–8224
86 Bennett, M. J., Schlegel, B. P., Jez, J. M., Penning, T. M. and Lewis, M. (1996) Biochemistry **35**, 10702–10711
87 Borhani, D. W., Harter, T. M. and Petrash, J. M. (1992) J. Biol. Chem. **267**, 24841–24847
88 Kubiseski, T. J., Hyndman, D. J., Morjana, N. A. and Flynn, T. G. (1992) J. Biol. Chem. **267**, 6510–6517
89 Bohren, K. M., Page, J. L., Shankar, R., Henry, S. P. and Gabbay, K. H. (1991) J. Biol. Chem. **266**, 24031–24037
90 Yamaoka, T., Matsuura, Y., Yamashita, K., Tanimoto, T. and Nishimura, C. (1992) Biochem. Biophys. Res. Commun. **183**, 327–333
91 Kubiseski, T. J. and Flynn, T. G. (1995) J. Biol. Chem. **270**, 16911–16917
92 Ehrig, T., Bohren, K. M., Prendergast, F. G. and Gabbay, K. H. (1994) Biochemistry **33**, 7157–7165
93 Petrash, J. M., Harter, T. M., Devine, C. S., Olins, P. O., Bhatnagar, A., Liu, S. Q. and Srivastava, S. K. (1992) J. Biol. Chem. **267**, 24833–24840
94 Bohren, K. M., Grimshaw, C. E. and Gabbay, K. H. (1992) J. Biol. Chem. **267**, 20965–20970
95 Barski, O. A., Gabbay, K. H. and Bohren, K. M. (1996) Biochemistry **35**, 14276–14280
96 Wilson, D. K., Tarle, I., Petrash, J. M. and Quiocho, F. A. (1993) Proc. Natl. Acad. Sci. U.S.A. **90**, 9847–9851
97 Harrison, D. H., Bohren, K. M., Ringe, D., Petsko, G. A. and Gabbay, K. H. (1994) Biochemistry **33**, 2011–2020
98 Li, J., Vrielink, A., Brick, P. and Blow, D. M. (1993) Biochemistry **32**, 11507–11515
99 Reardon, D. and Farber, G. K. (1995) FASEB J. **9**, 497–503
100 Evans, S. V. (1990) J. Mol. Graph. **12**, 134–138
101 Kraulis, P. J. (1991) J. Appl. Crystallogr. **24**, 946–950