Biochemical Pharmacology, Vol. 54, pp. 639–647, 1997.

ISSN 0006-2952/97/$17.00 + 0.00
PII S0006-2952(97)00092-0

COMMENTARY

A New Nomenclature for the Aldo-Keto Reductase Superfamily*

Joseph M. Jez,† T. Geoffrey Flynn‡ and Trevor M. Penning§‖

DEPARTMENTS OF †BIOCHEMISTRY & BIOPHYSICS, AND §PHARMACOLOGY, UNIVERSITY OF PENNSYLVANIA MEDICAL SCHOOL, PHILADELPHIA, PA 19104, U.S.A.; AND ‡DEPARTMENT OF BIOCHEMISTRY, QUEEN'S UNIVERSITY, ONTARIO, CANADA K7L 3N6

ABSTRACT. The aldo-keto reductases (AKRs) represent a growing oxidoreductase superfamily. Forty proteins have been identified and characterized as AKRs, and an additional fourteen genes may encode proteins related to the superfamily. Found in eukaryotes and prokaryotes, the AKRs metabolize a wide range of substrates, including aliphatic aldehydes, monosaccharides, steroids, prostaglandins, and xenobiotics. This broad substrate specificity has caused problems in naming these proteins. Enzymes capable of these reactions have been referred to as aldehyde reductase (ALR1), aldose reductase (ALR2), and carbonyl reductase (ALR3); however, ALR3 is not a member of the AKR superfamily. Also, some AKRs have multiple names based upon substrate specificity. For example, human 3α-hydroxysteroid dehydrogenase (3α-HSD) type I is also known as dihydrodiol dehydrogenase 4 and chlordecone reductase. To address these issues, we propose a new nomenclature system for the AKR superfamily based on amino acid sequence identities. Cluster analysis of the AKRs shows seven distinct families at the 40% amino acid identity level. The largest family (AKR1) contains the aldose reductases, aldehyde reductases, and HSDs. Other families include the prokaryotic AKRs, the plant chalcone reductases, the *Shaker* channels, and the ethoxyquin-inducible aflatoxin B_1 aldehyde reductase. At the level of 60% amino acid identity, subfamilies are discernible. For example, the AKR1 family includes five subfamilies: (A) aldehyde reductases (mammalian); (B) aldose reductases; (C) HSDs; (D) Δ^4-3-ketosteroid-5β-reductases; and (E) aldehyde reductases (plant). This cluster analysis forms the basis for our nomenclature system. Recommendations for naming an aldo-keto reductase include the root symbol "AKR," an Arabic number designating the family, a letter indicating the subfamily when multiple subfamilies exist, and an Arabic numeral representing the unique protein sequence. For example, human aldehyde reductase would be assigned as AKR1A1. Our nomenclature is both systematic and expandable, thereby allowing assignment of consistent designations for newly identified members of the superfamily. BIOCHEM PHARMACOL **54**;6:639–647, 1997. © 1997 Elsevier Science Inc.

KEY WORDS. aldo-keto reductase; nomenclature; aldehyde reductase; aldose reductase; hydroxysteroid dehydrogenase; multiple sequence alignment

The AKRs¶ represent a growing superfamily of NAD(P)(H)-dependent oxidoreductases [1, 2]. Proteins of the AKR superfamily are monomeric $(\alpha/\beta)_8$-barrel proteins, about 320 amino acids in length, which bind NAD(P)(H) without a Rossmann-fold motif [3–7]. Found in mammals, amphibians, plants, yeast, protozoa, and bacteria, the AKRs metabolize a range of substrates including aliphatic aldehydes, monosaccharides, steroids, prostaglandins, polycyclic aromatic hydrocarbons, and isoflavinoid phytoalexins. In addition, many mammalian AKRs are drug targets. Aldose reductase (EC 1.1.1.21) has been implicated in the development of diabetic retinopathy and neuropathy because it converts glucose to the hyperosmotic sugar sorbitol, and inhibitors for this enzyme are sought as potential treatments for these complications of diabetes [8]. Inhibitors of aldehyde reductase (EC 1.1.1.2), which metabolizes neurotransmitter aldehydes produced by monoamine oxidase, may have anti-depressant properties [9]. The HSDs that belong to this superfamily convert potent steroid hormones into their inactive metabolites and regulate the amount of hormone available for steroid receptors [10, 11]. Inhibitors of these enzymes may modulate steroid hormone action. Furthermore, AKRs that transform prostaglandins, e.g. 3α-HSD (EC 1.1.1.213) and prostaglandin F synthase (EC 1.1.1.188), are targets for nonsteroidal anti-inflammatory drugs (NSAIDs) [12, 13]. The rapid progress in identifying members of the superfamily has led to problems in naming these proteins, and the need for an acceptable nomenclature system is acute.

Substrate specificity formed the basis of previous naming

* A preliminary account of this paper was presented at the 8th International Workshop on the Enzymology and Molecular Biology of Carbonyl Metabolism held in Deadwood, SD, between June 29 and July 3, 1996.

‖ Corresponding author: Trevor M. Penning, Ph.D., Department of Pharmacology, University of Pennsylvania School of Medicine, 3620 Hamilton Walk, Philadelphia, PA 19104-6084. Tel. (215) 898-9445; FAX (215) 573-2236; E-mail: penning@pharm.med.upenn.edu

¶ *Abbreviations:* AKR, aldo-keto reductase; ALR1, aldehyde reductase; ALR2, aldose reductase; ALR3, carbonyl reductase; HSD, hydroxysteroid dehydrogenase; 3α-HSD, 3α-hydroxysteroid dehydrogenase; UTR, untranslated region; and DD, dihydrodiol dehydrogenase.

schemes for members of this superfamily. Turner and Flynn [14] originally proposed a classification system that included the three major enzymes involved in carbonyl metabolism. Aldehyde reductase, aldose reductase, and carbonyl reductase (EC 1.1.1.184) were designated ALR1, ALR2, and ALR3, respectively. However, carbonyl reductase is a member of the short-chain alcohol dehydrogenase/reductase superfamily [15], and the adoption of this nomenclature was not universal. Also, recent additions to the superfamily have outdated this system. Likewise, the broad substrate specificity of some AKR proteins has led to single members of the superfamily having multiple names. For example, human liver 3α-HSD type I [16] is also known as dihydrodiol dehydrogenase type 4 (EC 1.3.1.20) [17] and as chlordecone reductase (EC 1.1.1.255) [18].

To clarify the naming of AKRs, we have proposed a new nomenclature for the superfamily. Our system uses divisions within the superfamily based upon amino acid sequence identity and provides a straightforward method for assigning a consistent nomenclature. This nomenclature is both systematic and flexible, and can accommodate newly identified members of this growing superfamily.

METHOD OF SEQUENCE COMPILATION AND COMPARISON

Members of the AKR superfamily were identified by screening the GenBank (GB), the Protein Identification Resource (PIR), and Swiss-Prot (SP) databases with the amino acid sequence of rat liver 3α-HSD [19] using the BLAST program [20, 21]. Additional sequences were added from the literature.

The multiple sequence alignment (Fig. 1) used PILEUP from the GCG system [22]. To minimize the number of gaps introduced into the alignments, the gap and gap length penalties were set to 5.0 and 0.25, respectively. In addition, alignments were not weighted based on predicted or known secondary structure. The resulting dendrogram showing the relative percent amino acid identity among the AKRs was generated by PILEUP.

OVERVIEW OF THE SUPERFAMILY

To date, forty members of the AKR superfamily have been cloned and the encoded protein expressed or characterized in some fashion (Table 1). While the mammalian aldehyde reductases, aldose reductases, and HSDs represent the bulk of the superfamily, other AKRs have been identified in plants, yeast, and bacteria. Also, some members of the superfamily are not carbonyl oxidoreductases. For example, the Δ^4-3-ketosteroid-5β-reductases function as carbon—carbon double bond reductases; ρ-crystallin is primarily a lens structural protein; and the role of the *Shaker* K^+ channel β-subunit in channel inactivation is unclear [58]. Fourteen additional sequences encoding potential AKRs were not included in these alignments because the proteins encoded by these cDNAs remain unidentified or uncharacterized (Table 2).

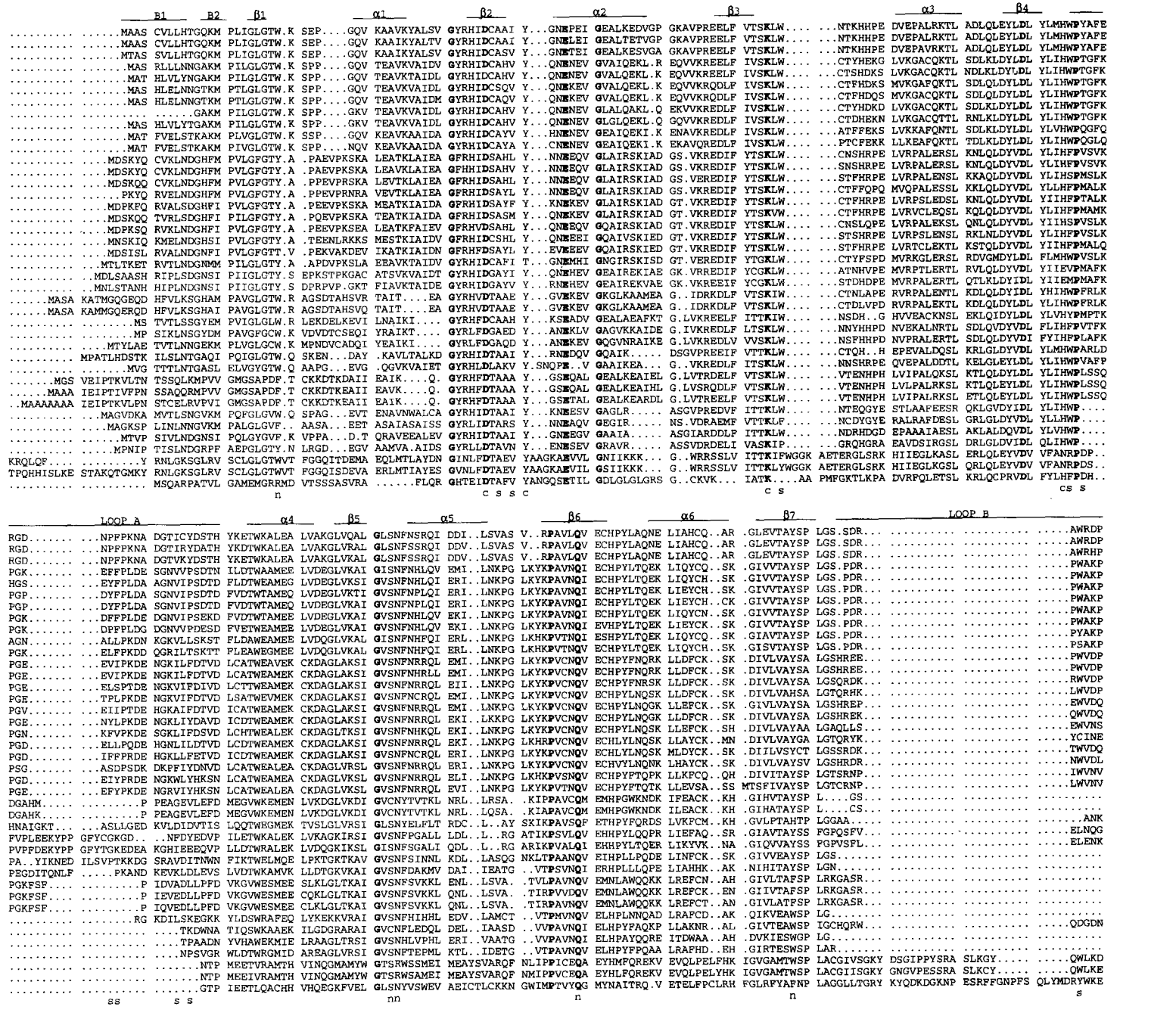
Multiple sequence comparison of the forty AKR amino acid sequences indicates that ten residues are invariant in the primary structures—G45, D50, E58, G62, K84, P119, G164, P186, Q190, and S271 (numbering corresponds to the sequence of rat liver 3α-HSD which was used for the alignment). Two of these residues (D50 and K84) along with Y55 and H117 (which are not strictly conserved) may have roles in catalysis. In other family members, D50, K84, Q190, and S271 are often involved in cofactor binding. When the secondary structure elements of rat liver 3α-HSD [6] are included above the alignment, other conserved residues are found in the β-strands, α-helices, and short loops of the barrel. This suggests that these residues may be important in maintaining the overall structure. Also, the size and composition of the eight β-strands and the eight α-helices forming the core of the barrel are generally conserved. The three loops found on the C-terminal side of the barrel display the greatest variation in both size and sequence, and could serve functional roles in ligand recognition. The need to maintain a common scaffold while altering loops to achieve ligand specificity is a feature of $(\alpha/\beta)_8$-barrel proteins [68].

Analysis using the percent amino acid identity of the AKRs showed that these proteins fall into seven distinct clusters within the superfamily (Fig. 2). The mammalian aldehyde reductases, aldose reductases, and HSDs, and the plant aldehyde reductases form the AKR1 cluster. The apple and yeast proteins of the AKR2 and AKR3 clusters metabolize monosaccharide substrates similar to those used by the aldehyde reductases and aldose reductases of the previous family. Cluster AKR4 consists of a polyketide reductase and the chalcone reductases; and cluster AKR5 is formed by the prokaryotic and protozoan AKRs. The most distantly related clusters are the *Shaker* K^+ channel β-subunits of AKR6 and the ethoxyquin-inducible aflatoxin B_1 aldehyde reductase of AKR7. The division of the superfamily into seven distinct clusters, or families, provides a rational basis for a new AKR nomenclature system.

THE NEW NOMENCLATURE

The system we propose is similar to the nomenclature for the cytochrome P450 superfamily [69], but, unlike that system, the nomenclature for the AKR superfamily uses amino acid sequence comparisons since detailed gene structures of most AKRs remain unknown at this time. The general format for new AKR names will be as follows: the root symbol "AKR" for Aldo-Keto Reductase; an Arabic number designating the family; a letter indicating the subfamily when multiple subfamilies exist; and an Arabic numeral representing the unique protein sequence. Under this system, the protein AKR1A1 would be the first AKR in family 1, subfamily A.

Delineation of families occurs at the 40% amino acid identity level. Members of an AKR family should have


β1 β2 β1 α1 β2 α2 β3 α3 β4
Hum_ALR
Por_ALR
Rat_ALR
Hum_ADR
Rab_ADR
Mou_ADR
Rat_ADR
Bov_ADR
Por_ADR
Mou_VDP
Mou_FR1
Hum_20aHSD
Hum_BABP
Hum_3aHSD2
Hum_ChlorRed
Rab_20aHSD
Mou_17bHSD
Bov_Pgs
Rat_20aHSD
Rat_3aHSD
Frog_Rho
Hum_3o5bRed
Rat_3o5bRed
Bar_ALR
Brgr_ALR
App_S6Pdh
Pic_XylRed
Klu_XylRed
Sac_GCY
Sporo_ALR
Alf_ChalRed
Soy_ChalRed
Gly_PKR
Lei_putRed
Ps_Mordh
Cb_25dkg
Cb_25dkg2
Bov_Shaker
Rat_Shaker
Rat_AFBRed
LOOP A α4 β5 α5 β6 α6 β7 LOOP B

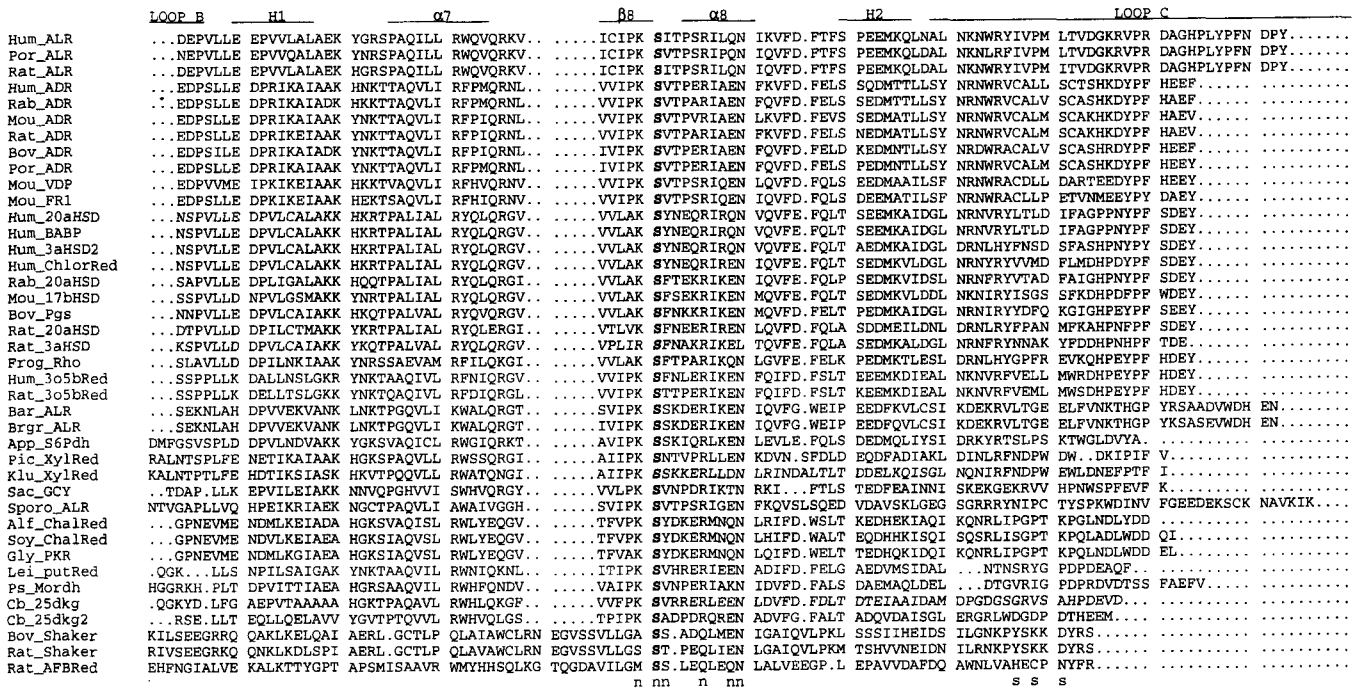


FIG. 1. Multiple sequence alignment of the AKR superfamily. The multiple sequence alignment was generated using PILEUP from the GCG program suite [22]. Above the alignment, the secondary structure of rat liver 3α-HSD is noted [6]. B1, B2, H1, and H2 are β-sheets and α-helices not involved in forming the core $(\alpha/\beta)_8$-barrel structure. Also, the residues corresponding to the three loops (loops A, B, and C) on the C-terminal side of the barrel are noted. Invariant residues are indicated in bold-face. Amino acids involved in catalysis (c), NAD(P)(H) binding (n), and substrate binding (s) are as indicated. Abbreviations for the proteins are as in Table 1. Reprinted with permission from Ref. 23. Copyright (1997) Plenum Publishing Corp.

TABLE 1. Member of the aldo-keto reductase superfamily

Abbreviation	Enzyme	Species	Accession	Ref.
Cb_25dkg	2,5-diketo-D-gulonic acid reductase (EC 1.1.1.125)	*Corynebacterium* sp.	GB: M12799	24
Cb2_25dkg	2,5-diketo-D-gulonic acid reductase (EC 1.1.1.125)	*Corynebacterium* sp.	SP: P15339	25
Lei_putRed	putative reductase	*Leishmania major*	SP: P22045	26
Ps_Mordh	morphine dehydrogenase (EC 1.1.1.218)	*Pseudomonas putida*	GB: M94775	27
Alf_ChalRed	chalcone reductase	*Medicago sativa*	PIR: S48851	28
Soy_ChalRed	chalcone reductase	*Glycine max*	SP: P26690	29
Gly_PKR	polyketide reductase	*Glcyrrhiza echinata*	GB: D83718	30
Hum_3o5bred	Δ^4-3-ketosteroid-5β-reductase (EC 1.3.99.6)	*Homo sapiens*—liver	PIR: S41120	31
Rat_3o5bred	Δ^4-3-ketosteroid-5β-reductase (EC 1.3.99.6)	*Rattus norvegicus*—liver	SP: P31210	32
Rat_3aHSD	3α-HSD (EC 1.1.1.213)	*Rattus norvegicus*—liver	SP: P23457	19
Hum_3aHSDII	3α-HSD type II (EC 1.1.1.213)	*Homo sapiens*—liver		16
Hum_BABP	bile acid binding protein dihydrodiol dehydrogenase 2 (EC 1.3.1.20) 3α-HSD (EC 1.1.1.213)	*Homo sapiens*—liver	PIR: S43843	33, 34
Hum_ChlorRed	chlordecone reductase (EC 1.1.1.125) dihydrodiol dehydrogenase 4 (EC 1.3.1.20) 3α-HSD type I (EC 1.1.1.213)	*Homo sapiens*—liver	PIR: S43844	18
Mou_17bHSD	17β-HSD (EC 1.1.1.62)	*Mus musculus*—liver	PIR: A56424	35
Hum_20aHSD	20α-HSD (EC 1.1.1.149) dihydrodiol dehydrogenase 1 (EC 1.3.1.20)	*Homo sapiens*—liver	PIR: A53436	33, 34
Rat_20aHSD	20α-HSD (EC 1.1.1.149)	*Rattus norvegicus*—ovary	PIR: S43842	36
Rab_20aHSD	20α-HSD (EC 1.1.1.149)	*Oryctolagus cuniculus*—ovary	PIR: A45366	11
Bov_Pgs	prostaglandin F synthase (EC 1.1.1.188)	*Bos taurus*—lung	GB: J03570	13
Frog_Rho	ρ-crystallin	*Rana catesbeiana*	GB: X87724	37
Bov_ADR	aldose reductase (EC 1.1.1.21)	*Bos taurus*—lens/testis	SP: P16116	38
Por_ADR	aldose reductase (EC 1.1.1.21)	*Sus scrofa*—lens	SP: P80276	39
Hum_ADR	aldose reductase (EC 1.1.1.21)	*Homo sapiens*—placenta	GB: J04795	1
Rab_ADR	aldose reductase (EC 1.1.1.21)	*Oryctolagus cuniculus*—kidney	SP: P15122	40
Rat_ADR	aldose reductase (EC 1.1.1.21)	*Rattus norvegicus*—lens	PIR: A60603	41
Mou_ADR	aldose reductase (EC 1.1.1.21)	*Mus musculus*—mouse	SP: 45376	42
Mou_VDP	androgen-dependent vas deferens protein	*Mus musculus*	GB: 81448	43
Mou_FR1	fibroblast growth factor induced protein	*Mus musculus*	SP: P45377	44
Hum_ALR	aldehyde reductase (EC 1.1.1.2)	*Homo sapiens*—liver	SP: 14550	45
Por_ALR	aldehyde reductase (EC 1.1.1.2)	*Sus scrofa*	GB: U46064	46
Rat_ALR	aldehyde reductase (EC 1.1.1.2)	*Rattus norvegicus*—liver	GB: D10853	47
Bar_ALR	aldehyde reductase (EC 1.1.1.2)	*Hordeum vulgare*	GB: Z48360	48
Brgr_ALR	aldehyde reductase (EC 1.1.1.2)	*Bromus inermis*	PIR: JQ2253	49
App_S6Pdh	sorbitol phosphate dehydrogenase	*Malus domestica*	SP: P28475	50
Pic_XylRed	xylose reductase	*Pichia stipitis*	SP: P31867	51
Klu_XylRed	xylose reductase	*Kluyveromyces lactis*	SP: P49378	52
Spor_ALR	aldehyde reductase (EC 1.1.1.2)	*Sporidiobolus salmonicolor*	GB: U26463	53
Sac_GCY	GCY protein	*Saccharomyces cerevisiae*	SP: P14065	54
Bov_Shaker	β2 subunit of *Shaker* K^+ channel	*Bos taurus*	GB: X70661	55
Rat_Shaker	β1 subunit of *Shaker* K^+ channel	*Rattus norvegicus*	GB: X70662	56
Rat_AFBred	aflatoxin B_1 aldehyde reductase	*Rattus norvegicus*—liver	GB: X74673	57

Accession numbers are from the GenBank (GB), the Protein Identification Resource (PIR), and the Swiss-Prot (SP) databases.

≤40% amino acid identity with any other family. At present, the seven families defined by our cluster analysis satisfies this criterion.

Within a given family, subfamilies may be defined by a ≥60% identity in amino acid sequence among subfamily members. By this definition, four of the seven AKR families include multiple subfamilies. For example, family AKR1 includes the following subfamilies: (A) mammalian aldehyde reductases; (B) mammalian aldose reductases; (C) the HSDs; (D) the Δ^4-3-ketosteroid-5β-reductases; and (E) plant aldehyde reductases. Numbering of the known members of each subfamily was assigned in an arbitrary fashion. For example, AKR1A1, AKR1A2, and AKR1A3 are the aldehyde reductases from human, pig, and rat, respectively. Any new additions to a subfamily will be numbered chronologically.

Allelic variation may occur between superfamily members. We propose that proteins with ≥97% amino acid sequence identity are alleles of the same gene unless: they have different enzyme activities; they are encoded by different cDNAs, usually evident by a distinct 3′-untranslated region (UTR); and they are derived from genes of different structure. While human dihydrodiol dehydrogenase 1 (DD1) and human dihydrodiol dehydrogenase 2 (DD2) are 98% identical in amino acid sequence and have 3′-UTRs that are 97% identical, the substrate specificity

TABLE 2. Fourteen potential members of the aldo-keto reductase superfamily

Enzyme	Species	Accession	Ref.
NRK-1-CDC-12 intergenic region	*Saccharomyces cerevisiae*	SP: P38715	59
YSW-1-RIB-7 intergenic region	*Saccharomyces cerevisiae*	SP: P38115	60
ACR1-YUH-7 intergenic region	*Saccharomyces cerevisiae*	SP: P47137	61
ASPU-DNIR intergenic region	*Escherichia coli*	SP: P30863	62
AKR-related gene	*Avena fatua*	GB: U21747	63
AKR-related from cosmid C01G5	*Caenorhabditis elegans*	GB: U50068	64
AKR-related from cosmid C35D10	*Caenorhabditis elegans*	GB: U21324	64
AKR-related gene	*Babesia bovis*	SP: P40690	65
igrA ORF	*Pseudomonas* sp.	SP: M37389	66
rat AKRs B, C, D, E, and F cDNAs	*Rattus norvegicus*		67

Accession numbers are from the GenBank (GB) and the Swiss-Prot (SP) databases.

and function of these proteins are quite different [33]. DD1 is predominantly a 20α-HSD while DD2 is the major bile acid binding protein in human liver. Based on these functional differences, we have assigned DD1 and DD2 as unique members of the AKR superfamily. However, it is not known with any certainty if these AKRs represent alleles of the same gene, and resolution of this issue will require genomic cloning and analysis of gene structure.

Finally, the designation for an AKR superfamily gene should be noted in italics to distinguish between the gene and the protein. For example, the gene *AKR1A1* encodes the protein AKR1A1.

NAMING THE ALDO-KETO REDUCTASES

At the 8th International Workshop on the Enzymology and Molecular Biology of Carbonyl Metabolism, the proposed nomenclature system was adopted. It was recommended that a World-Wide Web page should be established for the submission of new AKR sequences to our existing database (see below).

The new designations for the forty AKRs using the above criteria are indicated in Fig. 2. For historical reasons, the AKR1A subfamily represents the aldehyde reductases and the AKR1B subfamily represents the aldose reductases. To ease the adoption of the nomenclature system, we recommend that authors referencing members of the AKR superfamily use the old name with the new designation in parentheses until universal adoption of the system occurs—for example, human aldehyde reductase (AKR1A1).

Since our nomenclature system is based on amino acid sequence comparisons, we have not assigned designations to the fourteen potential AKRs listed in Table 2 because the proteins have not been expressed or characterized. We recommend that investigators assign a function to their AKR cDNA clones.

To facilitate the naming of newly discovered AKRs, a web-site has been established at http://pharme26.med.upenn.edu. This web-site includes the tables and figures of this paper, and will be updated as new AKRs are added to the superfamily. It also provides instructions on how to deposit new AKR sequences for inclusion into this nomenclature system. The procedure for the submission of a new sequence to obtain a systematic name is as follows:

(1) Since the proposed nomenclature system is protein-based, the newly identified AKR will require that the amino acid sequence has been obtained by either cDNA cloning or by direct methods. The protein encoded by a cDNA should be either overexpressed or purified from its natural source. Investigators should provide GenBank, Swiss-Prot, or PIR accession numbers.
(2) Upon submission of a complete protein sequence, it will be matched against the AKRs in the database and placed within the cluster analysis. The location of the sequence within the superfamily cluster analysis will determine its assigned designation. As needed, new families and subfamilies will be added to the existing system.
(3) The sequence and the assigned designation will be returned to the submitter, but the database will not be updated until the submission has been published. We encourage the submitter to use the new assignment in their publication. It is an investigator's responsibility to notify the web-site that the information submitted has been published and provide the appropriate citation.

SUMMARY

The need for a systematic and flexible nomenclature has arisen as new members of the AKR superfamily have been identified. The proposed system is similar to that used in the P450 superfamily but is based on amino acid sequence comparisons and not gene structure. Our system uses percent amino acid identity to define families and subfamilies within the larger AKR superfamily. We have also outlined a submission procedure for the assignment of names to newly cloned and characterized AKR proteins. As new AKRs are added to those already identified, this flexible nomenclature system will facilitate the easy incorporation of new proteins into the superfamily and the assignment of consistent designations to these proteins. Placement of new AKRs within the superfamily may ultimately lead to prediction of function.

This work was supported, in part, by NIH Grant DK47015 awarded to T.M.P.

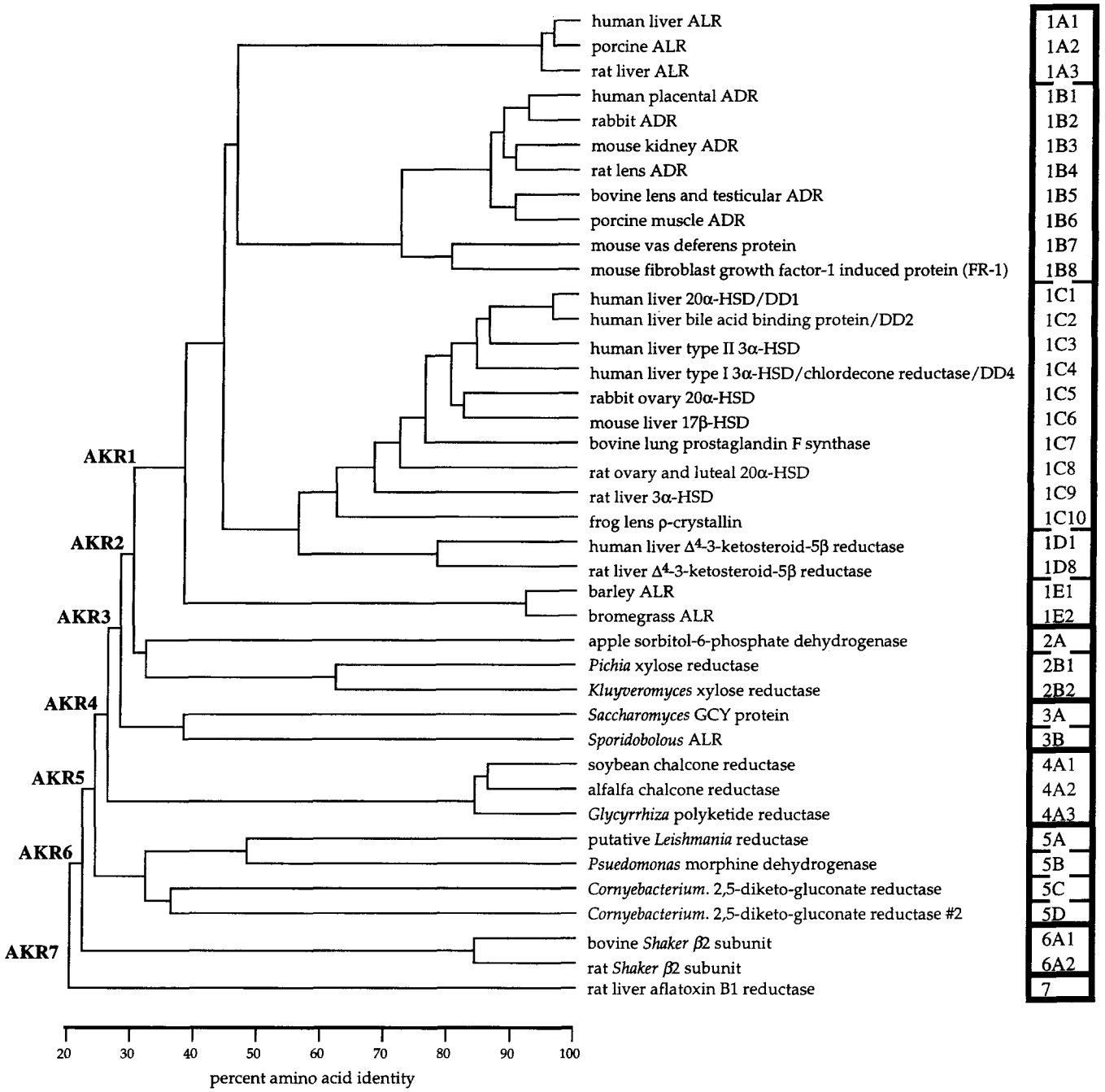


FIG. 2. Cluster analysis of the AKR superfamily. The dendrogram is based on the pairwise-sequence alignments performed by the GCG program PILEUP and indicates the percent amino acid identity among proteins of the AKR superfamily. Reprinted with permission from Ref. 23. Copyright (1997) Plenum Publishing Corp.

References

1. Bohren KM, Bullock B, Wermuth B and Gabbay KH, The aldo-keto reductase superfamily. *J Biol Chem* **264:** 9547–9551, 1989.
2. Bruce NC, Willey DL, Coulson AFW and Jeffrey J, Bacterial morphine dehydrogenase further defines a distinct superfamily of oxidoreductases with diverse functional activities. *Biochem J* **299:** 805–811, 1994.
3. Rondeau JM, Tete-Favier F, Podjarny A, Reymann JM, Barth P, Biellman JF and Moras D, Novel NADPH-binding domain revealed by the crystal structure of aldose reductase. *Nature* **355:** 469–472, 1992.
4. Wilson DK, Bohren KM, Gabbay KH and Quiocho FA, An unlikely sugar substrate site in the 1.65 Å structure of the human aldose reductase holoenzyme implicated in diabetic complications. *Science* **257:** 81–84, 1992.
5. Wilson DK, Nakano T, Petrash JM and Quiocho FA, 1.7 Å Structure of FR-1, a fibroblast growth factor-induced member of the aldo-keto reductase family, complexed with coenzyme and inhibitor. *Biochemistry* **34:** 14323–14330, 1995.
6. Hoog SS, Pawlowski JE, Alzari PM, Penning TM and Lewis M, Three-dimensional structure of rat liver 3α-hydroxy-steroid/dihydrodiol dehydrogenase: A member of the aldo-keto reductase superfamily. *Proc Natl Acad Sci USA* **91:** 2517–2521, 1994.

7. El-Kabbani I, Judge K, Ginell SL, Myles DAA, DeLucas LJ and Flynn TG, Structure of porcine aldehyde reductase holoenzyme. *Nat Struct Biol* **2:** 687–692, 1995.
8. Tomlinson DR, Stevens EJ and Diemel LT, Aldose reductase inhibitors and their potential for the treatment of diabetic complications. *Trends Pharmacol Sci* **15:** 293–297, 1994.
9. Turner AJ, Illingworth JA and Tipton KF, Simulation of biogenic amine metabolism in the brain. *Biochem J* **144:** 353–360, 1972.
10. Dorfman RI and Shipley RA, *Androgens*. John Wiley, New York, 1956.
11. Lacy WR, Washenick KJ, Cook RG and Dunbar BS, Molecular cloning and expression of an abundant rabbit ovarian protein with 20α-hydroxysteroid dehydrogenase activity. *Mol Endocrinol* **7:** 58–66, 1993.
12. Penning TM and Talalay P, Inhibition of a major NAD(P)-linked oxidoreductase from rat liver cytosol by steroidal and nonsteroidal anti-inflammatory agents and by prostaglandins. *Proc Natl Acad Sci USA* **80:** 4504–4508, 1982.
13. Watanabe K, Fujii Y, Nakayama K, Ohkubo H, Kuramitsu S, Kagamiyama H, Nakanishi S and Hayaishi O, Structural similarity of bovine lung prostaglandin F synthase to lens ε-crystallin of the European common frog. *Proc Natl Acad Sci USA* **85:** 11–15, 1988.
14. Turner AJ and Flynn TG, The nomenclature of aldehyde reductases. In: *Enzymology of Carbonyl Metabolism: Aldehyde Dehydrogenase and Aldo/Keto Reductase* (Eds. Weiner H and Wermuth B), pp. 401–402. Alan R. Liss, New York, 1982.
15. Jornvall H, Persson M, Krook M, Atrian S, Gonzalez-Duarte R, Jeffrey J and Ghosh D, Short-chain dehydrogenases/reductases (SDR). *Biochemistry* **34:** 6003–6013, 1995.
16. Khanna M, Qin KN, Wang RW and Cheng KC, Substrate specificity, gene structure, and tissue-specific distribution of multiple human 3α-hydroxysteroid dehydrogenases. *J Biol Chem* **270:** 20162–20168, 1995.
17. Deyashiki Y, Ogasawara A, Nakayama T, Nakanishi M, Miyabe Y, Sato K and Hara A, Molecular cloning of two human liver 3α-hydroxysteroid/dihydrodiol dehydrogenase isoenzymes that are identical with chlordecone reductase and bile-acid binder. *Biochem J* **299:** 545–552, 1994.
18. Winters CJ, Molowa DT and Guzelian PS, Isolation and characterization of cloned cDNAs encoding human liver chlordecone reductase. *Biochemistry* **29:** 1080–1087, 1990.
19. Pawlowski JE, Huizinga M and Penning TM, Cloning and sequencing of the cDNA for rat liver 3α-hydroxysteroid/dihydrodiol dehydrogenase. *J Biol Chem* **266:** 8820–8825, 1991.
20. Altschul SF, Gish W, Miller W, Myers EW and Lipman DJ, Basic local alignment search tool. *J Mol Biol* **215:** 403–410, 1990.
21. Gish W and States DJ, Identification of protein coding regions by database similarity search. *Nat Gene* **3:** 266–272, 1993.
22. Devereux J, Haeberil P and Smithies O, A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res* **12:** 387–395, 1984.
23. Jez JM, Flynn TG and Penning TM, A proposed nomenclature for the aldo-keto reductase superfamily. In: *Enzymology and Molecular Biology of Carbonyl Metabolism 6, Advances in Experimental Medicine and Biology* (Eds. Weiner H, Lindahl R, Crabb DW and Flynn TG), Vol. 414, pp. 579–590. Plenum Press, New York, 1997.
24. Anderson S, Berman-Marks C, Lazarus R, Miller J, Stafford K, Seymour J, Light D, Rastetter W and Estell D, Production of 2-keto-L-gulonate, an intermediate in L-ascorbate synthesis, by a genetically modified *Erwinia herbicola*. *Science* **230:** 144–149, 1985.
25. Grindley JG, Payton MA, De Pol H and Hardy KG, Conversion of glucose to 2-keto-L-gulonate, an intermediate in L-ascorbate synthesis, by a recombinant strain of *Erwinia citreus*. Appl Environ Microbiol **54:** 1770–1775, 1988.
26. Samaras N and Spithill TW, The developmentally regulated P100/11E gene of *Leishmania major* shows homology to a superfamily of reductase genes. *J Biol Chem* **264:** 4251–4254, 1989.
27. Willey DL, Caswell DA, Lowe CR and Bruce NC, Nucleotide sequence and over-expression of morphine dehydrogenase, a plasmid-encoded gene from *Pseudomonas putida* M10. *Biochem J* **290:** 539–544, 1993.
28. Sallaud C, el Turk C, Bigerre L, Sevin H, Welle R and Esnault R, Nucleotide sequence of three chalcone reductase genes from alfalfa. *Plant Physiol* **108:** 869–870, 1995.
29. Welle R, Schröder G, Schiltz E, Grisebach H and Schröder J, Induced plant responses to pathogen attack. Analysis and heterologous expression of the key enzyme in the biosynthesis of phytoalexins in soybean (*Glycine max* L. Merr. cv. Harosoy 63). *Eur J Biochem* **196:** 423–430, 1991.
30. Akashi T, Furuno T, Futami K, Honda M, Takahashi T, Welle R and Ayabe S, A cDNA for polyketide reductase (accession No. D83718) that catalyzes the formation of 6′-deoxychalcone from cultured *Glycyrrhiza echinata* L. cells (PGR 96-023). *Plant Physiol* **111:** 347, 1996.
31. Kondo K, Kai M, Setoguchi Y, Eggertsen G, Sjoblom P, Setoguchi T, Okuda K and Bjorkhem I, Cloning and expression of cDNA of human Δ^4-3-oxosteroid-5β-reductase and substrate specificity of the expressed enzyme. *Eur J Biochem* **219:** 357–363, 1994.
32. Onishi Y, Noshiro M, Shimosato T and Okuda K, Molecular cloning and sequence analysis of cDNA encoding Δ^4-3-ketosteroid-5β-reductase of rat liver. *FEBS Lett* **283:** 215–218, 1994.
33. Hara A, Matsuura K, Tamada Y, Sato K, Miyabe Y, Deyashiki Y and Ishida N, Relationship of human liver dihydrodiol dehydrogenases to hepatic bile-acid binding protein and an oxidoreductase of human colon cells. *Biochem J* **313:** 373–376, 1996.
34. Stolz A, Hammond L, Lou H, Takikawa H, Ronk M and Shively JE, cDNA cloning and expression of the human bile-acid binding protein. *J Biol Chem* **268:** 10448–10457, 1993.
35. Deyashiki Y, Ohshima K, Nakanishi M, Sato K, Matsuura K and Hara A, Molecular cloning and characterization of mouse estradiol 17β-dehydrogenase (A-specific), a member of the aldo-keto reductase family. *J Biol Chem* **270:** 10461–10467, 1995.
36. Miura R, Shiota K, Noda K, Yagi S, Ogawa T and Takahashi M, Molecular cloning of cDNA for rat ovarian 20α-hydroxysteroid dehydrogenase (HSD1). *Biochem J* **299:** 561–567, 1994.
37. Tomarev SI, Zinovieva RD, Dolgilevich SM, Luchin SV, Krayev AS, Skryabin KG and Gause GG, A novel type of crystallin in the frog eye lens. *FEBS Lett* **171:** 297–301, 1984.
38. Petrash JM and Favello AD, Isolation and characterization of cDNA clones encoding aldose reductase. *Exp Eye Res* **8:** 1021–1027, 1989.
39. Jaquinod M, Potier N, Klarskov K, Reymann JM, Sorokine O, Kieffer S, Barth P, Andriantomanga V, Biellman JF and Van Dorsselaes A, Sequence of pig lens aldose reductase and electrospray mass spectroscopy of non-covalent and covalent complexes. *Eur J Biochem* **218:** 893–903, 1993.
40. Garcia-Perez A, Martin B, Murphy HR, Uchida S, Murer H, Cowley BD, Handler JS and Burg MB, Molecular cloning of cDNA coding for kidney aldose reductase. *J Biol Chem* **264:** 16815–16821, 1989.
41. Carper D, Nishimura C, Shinohara T, Dietzchold B, Wistow G, Craft C, Kador P and Kinoshita JH, Aldose reductase and

ρ-crystallin belong to the same protein superfamily as aldehyde reductase. *FEBS Lett* **220:** 209–213, 1987.
42. Gui T, Tanimoto T, Kokai Y and Nishimura C, Presence of a closely related subgroup in the aldo-keto reductase superfamily of the mouse. *Eur J Biochem* **227:** 448–453, 1995.
43. Pailhoux EA, Martinez A, Veyssiere GM and Jean CG, Androgen-dependent protein from mouse vas deferens. *J Biol Chem* **265:** 19932–19936, 1990.
44. Donohue PJ, Alberts GF, Hampton BS and Winkles JA, A delayed-early gene activated by fibroblast growth factor-1 encodes a protein related to aldose reductase. *J Biol Chem* **269:** 8604–8609, 1994.
45. Wermuth B, Omar A, Forster A, di Francesco G, Wolf M, von Wartbuth JP, Bullock B and Gabbay KH, Primary structure of aldehyde reductase from human liver. *Prog Clin Biol Res* **232:** 297–307, 1987.
46. Flynn TG, Green NC, Bhatia MB and El-Kabbani O, Structure and mechanism of aldehyde reductase. *Adv Exp Med Biol* **372:** 193–201, 1995.
47. Takahashi M, Fujii J, Teshima T, Suzuki K, Shiba T and Taniguchi N, Identity of a major 3-deoxyglucosone-reducing enzyme with aldehyde reductase in rat liver established by amino acid sequencing and cDNA expression. *Gene* **127:** 249–253, 1993.
48. Bartels D, Engelhardt K, Roncarati R, Schneider K, Rutter M and Salamini F, An ABA and GA modulated gene expressed in the barley embryo encodes an aldose reductase related protein. *EMBO J* **10:** 1037–1043, 1991.
49. Lee SP and Chen THH, Molecular cloning of abscisic acid-responsive mRNAs expressed during the induction of freezing tolerance in bromegrass (*Bromus inermis* Leyss) suspension culture. *Plant Physiol* **101:** 1089–1096, 1993.
50. Kanayama Y, Moni H, Imaseki H and Yamaki S, Nucleotide sequence of a cDNA encoding NADP-sorbitol-6-phosphate dehydrogenase from apple. *Plant Physiol* **100:** 1607–1608, 1992.
51. Amore R, Kötter P, Küster C, Ciriacy M and Hollenberg CP, Cloning and expression in *Saccharomyces cerevisiae* of the NAD(P)H-dependent xylose reductase-encoding gene (*XYL1*) from the xylose-assimilating yeast *Pichia stipitis*. *Gene* **109:** 89–97, 1991.
52. Billard P, Menart S, Fleer R and Bolotin-Fukuhara M, Isolation and characterization of the gene encoding xylose reductase from *Kluyveromyces lactis*. *Gene* **162:** 93–97, 1995.
53. Kita K, Matsuzaki K, Hashimoto T, Yanase H, Kato N, Chung MCM, Kataoka M and Shimizu S, Cloning of the aldehyde reductase gene from a red yeast, *Sporobolomyces salmonicolor*, and characterization of the gene and its product. *Appl Environ Microbiol* **62:** 2303–2310, 1996.
54. Oechsner U, Magdolen V and Bandlow W, A nuclear yeast gene (GCY) encodes a polypeptide with high homology to a vertebrate lens protein. *FEBS Lett* **238:** 123–128, 1988.
55. Scott VES, Rettig J, Parcej DN, Keen JN, Findlay JBC, Pongs O and Dolly JO, Primary structure of a β-subunit of α-dendrotoxin-sensitive K^+ channels from bovine brain. *Proc Natl Acad Sci USA* **91:** 1637–1641, 1994.
56. Rettig J, Heinemann SH, Wunder F, Lorra C, Parcej CN, Dolly JO and Pongs O, Inactivation properties of voltage-gated K^+ channels altered by presence of subunit. *Nature* **369:** 289–294, 1994.
57. Ellis EM, Judah DJ, Neal GE and Hayes JD, An ethoxyquin-inducible aldehyde reductase from rat liver that metabolizes aflatoxin B_1 defines a subfamily of aldo-keto reductases. *Proc Natl Acad Sci USA* **90:** 10350–10354, 1993.
58. McCormack T and McCormack K, *Shaker* K^+ channel β-subunits belong to an NAD(P)H-dependent oxidoreductase superfamily. *Cell* **79:** 1133–1135, 1994.
59. Johnston M, Andrews S, Brinkman R, Cooper J, Ding H, Dover J, Du Z, Favello A, Fulton L, Gattung S, Geisel C, Kirsten J, Kucaba T, Hillier L, Jier M, Johnston L, Langston Y, Latreille P, Louis EJ, Macri C, Mardis E, Menezes S, Mouser L, Nhan M, Rifkin L, Riles L, St. Peter H, Trevaskis E, Vaughan K, Vignati D, Wilcox L, Wohldman P, Waterston R, Wilson P and Vaudin M, Complete nucleotide sequence of *Saccharomyces cerevisiae* chromosome VIII. *Science* **265:** 2077–2082, 1994.
60. Martinez-Soriano JP, Wong WM, Van Ryk DI and Nazar RN, A widely distributed 'CAT' family of repetitive DNA sequences. *J Mol Biol* **217:** 629–635, 1991.
61. Ramezani M, Kirchrath L and Hollenberg CP, GenBank submission, 1995.
62. Nishimura K, Komine Y, Miyamoto K, Kitabatake M, Mathunga F, Hisano T, Miki T and Inokuchi H, GenBank submission, 1992.
63. Li B and Foley ME, Cloning and characterization of differentially expressed genes in imbibed dormant and afterripened *Avena fatua* embryos. *Plant Mol Biol* **29:** 823–831, 1995.
64. Wilson R, Ainscough R, Anderson K, Baynes C, Berks M, Bonfield J, Burton J, Connell M, Copsey T, Cooper J, Coulson A, Craxton M, Dear S, Du Z, Durbin R, Favello A, Fraser A, Fulton L, Gardner A, Green P, Hawkins T, Hillier L, Jier M, Johnston L, Jones M, Kershaw J, Kirsten J, Laisster N, Latreille P, Lightning J, Lloyd C, Mortimore B, O'Callaghan M, Parsons J, Percy C, Rifken L, Roopra A, Saunders D, Shownkeen R, Sims M, Smaldon N, Smith A, Smith M, Sonnhammer E, Staden R, Sulston J, Thierry-Mieg J, Thomas K, Vaudin M, Vaughan K, Waterston R, Watson A, Weinstock L, Wilkinson-Sproat J and Wohldman P, 2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*. *Nature* **368:** 32–38, 1994.
65. Dalrymple BP, Peters JM and Vuocolo T, Characterization of genes encoding two novel members of the aldo-keto reductase superfamily. *Biochem Int* **28:** 651–657, 1992.
66. Fitzgibbon JE and Braymer HD, Cloning of a gene from *Pseudomonas* sp. strain PG2982 conferring increased glyphosate resistance. *Appl Environ Microbiol* **56:** 3382–3388, 1990.
67. Qin KN and Cheng KC, Structure and tissue-specific expression of the aldo-keto reductase superfamily. *Biochemistry* **33:** 3223–3228, 1994.
68. Farber GK and Petsko GA, The evolution of α/β-barrel enzymes. *Trends Biochem Sci* **15:** 228–234, 1990.
69. Nelson DR, Kamataki T, Waxman DJ, Guengerich FP, Estabrook RW, Feyereisen R, Gonzalez FJ, Coon MJ, Gunsalus IC, Gotoh O, Okuda K and Nebert DW, The P450 superfamily: Update on new sequences, gene mapping, accession numbers, early trivial names of enzymes, and nomenclature. *DNA Cell Biol* **12:** 1–51, 1993.