Biochemistry

 Volume 41, Number 28 July 16, 2002

Articles

The Structure of Apo and Holo Forms of Xylose Reductase, a Dimeric Aldo-Keto Reductase from *Candida tenuis*[†,‡]

Kathryn L. Kavanagh,[#] Mario Klimacek,[§,⊥] Bernd Nidetzky,[§,⊥] and David K. Wilson*[,#]

Section of Molecular and Cellular Biology, University of California, Davis, California 95616, and Division of Biochemical Engineering, Institute of Food Technology, University of Agricultural Sciences Vienna, Muthgasse 18, A-1190 Vienna, Austria

Received March 7, 2002; Revised Manuscript Received May 22, 2002

ABSTRACT: Xylose reductase is a homodimeric oxidoreductase dependent on NADPH or NADH and belongs to the largely monomeric aldo-keto reductase superfamily of proteins. It catalyzes the first step in the assimilation of xylose, an aldose found to be a major constituent monosaccharide of renewable plant hemicellulosic material, into yeast metabolic pathways. It does this by reducing open chain xylose to xylitol, which is reoxidized to xylulose by xylitol dehydrogenase and metabolically integrated via the pentose phosphate pathway. No structure has yet been determined for a xylose reductase, a dimeric aldo-keto reductase or a family 2 aldo-keto reductase. The structures of the *Candida tenuis* xylose reductase apo- and holoenzyme, which crystallize in spacegroup *C*2 with different unit cells, have been determined to 2.2 Å resolution and an R-factor of 17.9 and 20.8%, respectively. Residues responsible for mediating the novel dimeric interface include Asp-178, Arg-181, Lys-202, Phe-206, Trp-313, and Pro-319. Alignments with other superfamily members indicate that these interactions are conserved in other dimeric xylose reductases but not throughout the remainder of the oligomeric aldo-keto reductases, predicting alternate modes of oligomerization for other families. An arrangement of side chains in a catalytic triad shows that Tyr-52 has a conserved function as a general acid. The loop that folds over the NAD(P)H cosubstrate is disordered in the apo form but becomes ordered upon cosubstrate binding. A slow conformational isomerization of this loop probably accounts for the observed rate-limiting step involving release of cosubstrate. Xylose binding (K_m = 87 mM) is mediated by interactions with a binding pocket that is more polar than a typical aldo-keto reductase. Modeling of xylose into the active site of the holoenzyme using ordered waters as a guide for sugar hydroxyls suggests a convincing mode of substrate binding.

Candida tenuis xylose reductase (XR; EC 1.1.1.21)[1] is a member of the aldo-keto reductase (AKR) superfamily of enzymes which comprise well over 100 different proteins (*1*). These proteins have been found in a wide variety of organisms but in many cases do not fulfill an obvious physiological function. As a member of the structurally

[†] This work was supported by a grant from the National Institutes of Health to D.K.W. and the Keck Foundation. K.L.K. is supported by a grant from the University of California Systemwide Biotechnology Research Program, proposal number 2001-07. Further support to B.N. was provided by the Austrian Science Foundation (Grant P-12569-MOB). The data collection facilities at Stanford Synchrotron Radiation Laboratory are funded by the U.S. Department of Energy and by the National Institutes of Health.

[‡] Accession numbers: the coordinates for apo and holo XR have been deposited into the Protein Data Bank under the ID codes 1JEZ and 1K8C, respectively.

* To whom correspondence should be addressed. Section of Molecular Biology, One Shields Ave., University of California, Davis, CA 95616, Phone: (530) 752-1136; fax: (530) 752-3085; e-mail: dave@alanine.ucdavis.edu.

[#] University of California.

[§] University of Agricultural Sciences Vienna.

[⊥] Current address: Institute of Biotechnology, Graz University of Technology, Petersgasse 12, A-8010 Graz, Austria.

10.1021/bi025786n CCC: $22.00
Published on Web 06/19/2002

uncharacterized AKR family 2, which consists of xylose and mannose reductases, XR catalyzes the reversible NAD(P)H-dependent reduction of D-xylose to xylitol. This reaction is the first step in assimilation of xylose into the glycolytic pathway (*2*). The physiological role of XR is therefore clearly defined. Since xylose is known to be among the most abundant sugars in biomass (*3*), there is considerable interest in improving its metabolic utilization. The structure of XR provides information that will lead to engineering of an enzyme with improved substrate and cosubstrate binding properties. Enhanced efficiency of fermentation could lead to the conversion of agricultural byproducts and waste products into useful compounds such as ethanol (*4*) and the sweetener xylitol (*5*).

The majority of the known AKRs are monomeric, but XR functions as a noncooperative, tightly associated dimer composed of 322-amino acid, 36-kDa subunits. The enzyme has been shown to follow an ordered bi-bi mechanism with the cofactor binding first and the carbonyl-containing substrate second. Although sequence alignments indicated the active site is very hydrophobic, subsequent kinetic experiments show that it is adequate for activity with physiological concentrations of xylose (K_m = 87 mM) (*6*). Yeasts lacking a functional form of the enzyme, however, are unable to grow on xylose as a carbon source. Although many AKRs have been shown to accommodate sugars as substrates, a clear structural understanding of sugar binding to any of these has not been reached, probably due to their low affinities. How these enzymes bind and achieve any specificity for polar substrates is therefore vague.

In general, the AKRs catalyze the NADPH-dependent reduction of carbonyl-containing substrates to their corresponding alcohols. In rarer cases, as in some XRs (including *C. tenuis* XR) and several other AKRs, the enzyme can also accommodate NADH as a cosubstrate. Although the reaction is reversible, the majority of the AKRs favor the direction in which the carbonyl substrate is reduced. There is usually little substrate specificity among these enzymes. Hydrophobic compounds are generally preferred over hydrophilic compounds, however, by several orders of magnitude with respect to catalytic efficiency. Organisms expressing multiple AKRs often exhibit overlapping substrate specificity among those enzymes. While the majority of the AKRs have no well-defined role assigned to them, many are believed to act as general detoxification catalysts by reducing reactive carbonyl-containing compounds (*7*). It has also been proposed that they function in osmotic regulation by controlling levels of intracellular polyols (*8*). Clinical interest in the AKRs stems primarily from the ability of human aldose reductase to reduce the open chain form of glucose to sorbitol, a reaction that becomes physiologically significant when blood glucose levels are elevated in diabetic patients (*9*). Compounds that inhibit human aldose reductase have been studied for more than two decades to develop drugs that could potentially combat these complications.

The AKR superfamily of enzymes has been the subject of a number of different structural studies. Initially, these were directed at understanding the mechanism and inhibition of aldose reductase. The number of proteins confirmed to be members of the family has expanded 10-fold from the dozen known at the beginning of the 1990s. Correspondingly, the number that has been structurally characterized has grown to eight, including members of families 1, 3, 5, and 6 of the 12 that are known. In all of these cases, they bind cosubstrate similarly by adopting a distinctive, non-Rossman nucleotide binding motif, the $(\beta/\alpha)_8$ barrel. The current study of XR explores the structural divergence of family 2 and its novel dimeric nature. It also details substrate and cosubstrate binding and supports a common catalytic mechanism for all of the AKRs.

EXPERIMENTAL PROCEDURES

Expression, Purification, and Molecular Mass Determination. The expression vector pBEAct.1i was used for the heterologous production of xylose reductase from *C. tenuis* CBS 4435 and has been described previously (*10*). Competent *Escherichia coli* BL21(DE3) cells were transformed with this construct, and the resulting strain was grown at 25 °C using recently reported conditions for cultivation, induction, and purification of recombinant protein (*11*).

Column sizing experiments were carried out on an Äktaexplorer 100 system (Amersham Pharmacia Biotech) using a Superdex 75 HR 10/30 column. The column was equilibrated with 20 mM Tris/HCl buffer, pH 7.0, containing 100 mM NaCl. Purified XR was diluted in the same buffer (1 mL; 0.3 mg/mL) and loaded onto the column. Elution was carried out at a flow rate of 0.5 mL/min. Protein elution was detected by absorbance at 280 nm. Calibration for molecular mass determination was done by using the following protein standards (Pharmacia): BSA (66 kDa); ovalbumin (45 kDa); carbonic anhydrase (29 kDa); cytochrome *c* (12.3 kDa).

Dynamic light scattering experiments were performed on a Protein Solutions DynaPro 99 molecular sizing instrument. The protein concentration was approximately 1 mg/mL in a buffer containing 10 mM Hepes, pH 7.4.

Crystallization and X-ray Data Collection. Crystallization of the apo form of xylose reductase was achieved by concentrating the purified protein to 16 mg/mL and changing the buffer to 10 mM Hepes, pH 7.4, for use in hanging drop vapor diffusion experiments. Crystals were obtained in several different independent conditions, but the best were grown by suspending a 2 μL drop containing 8 mg/mL protein, 15% w/v poly(ethylene glycol) 4000, 100 mM ammonium acetate, 50 mM sodium citrate, pH 5.6, over a well containing 30 % w/v poly(ethylene glycol) 4000, 200 mM ammonium acetate, 100 mM sodium citrate, pH 5.6. Irregularly shaped crystals measuring approximately 0.3 × 0.1 × 0.1 mm appeared after several days and were used for diffraction experiments.

Crystals were flash cooled in a buffer containing 25% v/v ethylene glycol and 75% v/v well solution. Diffraction intensities were collected on an R-AXIS IV imaging plate detector, and the program DENZO (*12*) was used to determine lattice constants of a = 110.84 Å, b = 63.35 Å, c = 103.16 Å, β = 111.28°. The lattice and systematic extinctions indicated the spacegroup was *C*2. Calculation of the Matthews' constant (V_m) of 2.34 Å^3/Da implied two

[1] Abbreviations: hAR, human aldose reductase; AKR, aldo-keto reductase; XR, *Candida tenuis* xylose reductase; 3αHSD, 3α-hydroxysteroid dehydrogenase; Kvβ2, voltage-dependent K+ channel β subunit; TIM, triosephosphate isomerase.

Table 1: Summary of XR Data Collection, Refinement, and Models[a]

	Apo-XR	Holo-XR
Data Collection	R-AXIS II /rotating anode	MAR 345/ SSRL 7-1
space group	*C*2	*C*2
unit cell	$a = 110.65$ Å	$a = 182.20$ Å
	$b = 63.45$ Å	$b = 127.81$ Å
	$c = 103.05$ Å	$c = 80.25$ Å
	$\beta = 111.239°$	$\beta = 90.366°$
monomers per asymmetric unit	2	4
resolution range (Å)	100−2.2	30−2.2
no. of observations/reflections	109,237/33,457	325,247/93,492
R_{merge} (overall/high-resolution shell)	0.089/0.406	0.074/0.374
completeness (overall/high-resolution shell) (%)	98.8/97.3	99.5 /98.7
mosaicity (°)	0.58	0.51
$I/\sigma(I)$ (overall/high-resolution shell)	12.8/3.1	17.2/3.7
Model		
protein atoms	4,770	10,111
$NADP^+$ atoms		192
water molecules	407	734
overall temperature factors (Å^2)		
monomer A	27.68	30.93
monomer B	27.82	22.89
monomer C		27.00
monomer D		32.72
Refinement		
reflections used ($I > 0$)	30,391	92,623
R_{cryst}	0.179	0.208
R_{free}	0.233	0.237
rms deviation from ideal bond length (Å)	0.009	0.009
rms deviation from ideal bond angle (°)	1.35	1.44

[a] High-resolution shell for apo-XR is 2.28−2.20 Å, for holo-XR, 2.24−2.20 Å.

protein monomers in the asymmetric unit (*13*). Data to 2.2 Å were integrated and used for structure determination and refinement (Table 1).

Crystals of holo XR were grown by hanging drop vapor diffusion by mixing equal volumes of protein solution containing 16 mg/mL protein, 5 mM NADPH, 10 mM Hepes, pH 7.4, and well solution containing 32% w/v monomethyl ether poly(ethylene glycol) 5000, 300 mM ammonium sulfate, 100 mM sodium citrate, pH 5.6 and suspending over a reservoir of well solution. Crystals were flash cooled in a buffer of 25% v/v ethylene glycol and 75% v/v well solution. Diffraction intensities were collected at Stanford Synchrotron Radiation Laboratory beamline 7-1 on a MAR 345 image plate detector. The program DENZO was used to determine the lattice constants of $a = 182.20$ Å, $b = 127.81$ Å, $c = 80.25$ Å, $\beta = 90.37°$ and indicated that the spacegroup was *C*2. Data to 2.2 Å were processed using the programs DENZO and SCALEPACK (*12*). Calculation of a Matthews' constant (V_m) of 3.24 Å^3/Da implied four protein monomers in the asymmetric unit.

Structure Determination and Refinement. Molecular replacement as implemented by the program EPMR (*14*) was used to determine the structure of the apo form of xylose reductase. The search model was a modified human aldose reductase (*15*) with the waters and the NADPH cofactor removed and residues that were not identical to those in xylose reductase mutated to alanine. An initial solution using diffraction data between 20 and 5 Å was found for two monomers that yielded a correlation coefficient of 0.56 and an initial R_{cryst} of 0.44 within this resolution range. Before refinement commenced, 5% of the data were flagged for calculation of R_{free}. Side chains were added to the initial structure followed by manual fitting and addition of ordered water molecules using the programs CHAIN (*16*) and O (*17*). Iterations of manual refitting were followed by crystallographic refinement against reflections where $I > 0$ using the program CNS (*18*). Statistics for the final structure appear in Table 1.

The structure of holo XR was also determined using EPMR (*14*) with the apo XR structure, stripped of ordered water molecules, used as a search model. The initial solution for two dimers per asymmetric unit using data in the resolution range from 20 to 4.0 Å yielded a correlation coefficient of 0.59 and R_{cryst} of 0.39. Again, 5% of data were flagged for R_{free}, and after rigid body refinement, maps were calculated in which electron density for the NADPH was obvious. Inclusion of the cofactor in the model reduced R_{cryst} to 0.28 and R_{free} to 0.33. Loops containing residues 218−238 were built and ordered water molecules were added to the model. Iterations of manual fitting and refinement using the programs O and CNS resulted in the final structure for which statistics appear in Table 1.

RESULTS AND DISCUSSION

Overall Structure. The models of XR derived from refinement against 2.2 Å data for the apo and holo forms (Table 1) show that the protein folds into a $(\beta/\alpha)_8$ barrel also known as a triose phosphate isomerase or TIM barrel (Figure 1). The $(\beta/\alpha)_8$ barrel has emerged as the most common fold seen in enzymes, regardless of function (*19*). The structure contained within the asymmetric unit of the apo structure consists of two monomers related by a noncrystallographic 2-fold. The holo form packs differently, and four monomers are found in the asymmetric unit. The overall dimensions of the dimer in both cases are approximately 90 × 60 × 60 Å (Figure 1A).

Apo monomer A consists of 300 amino acids[2] out of the 322 predicted from sequence data. The first three amino

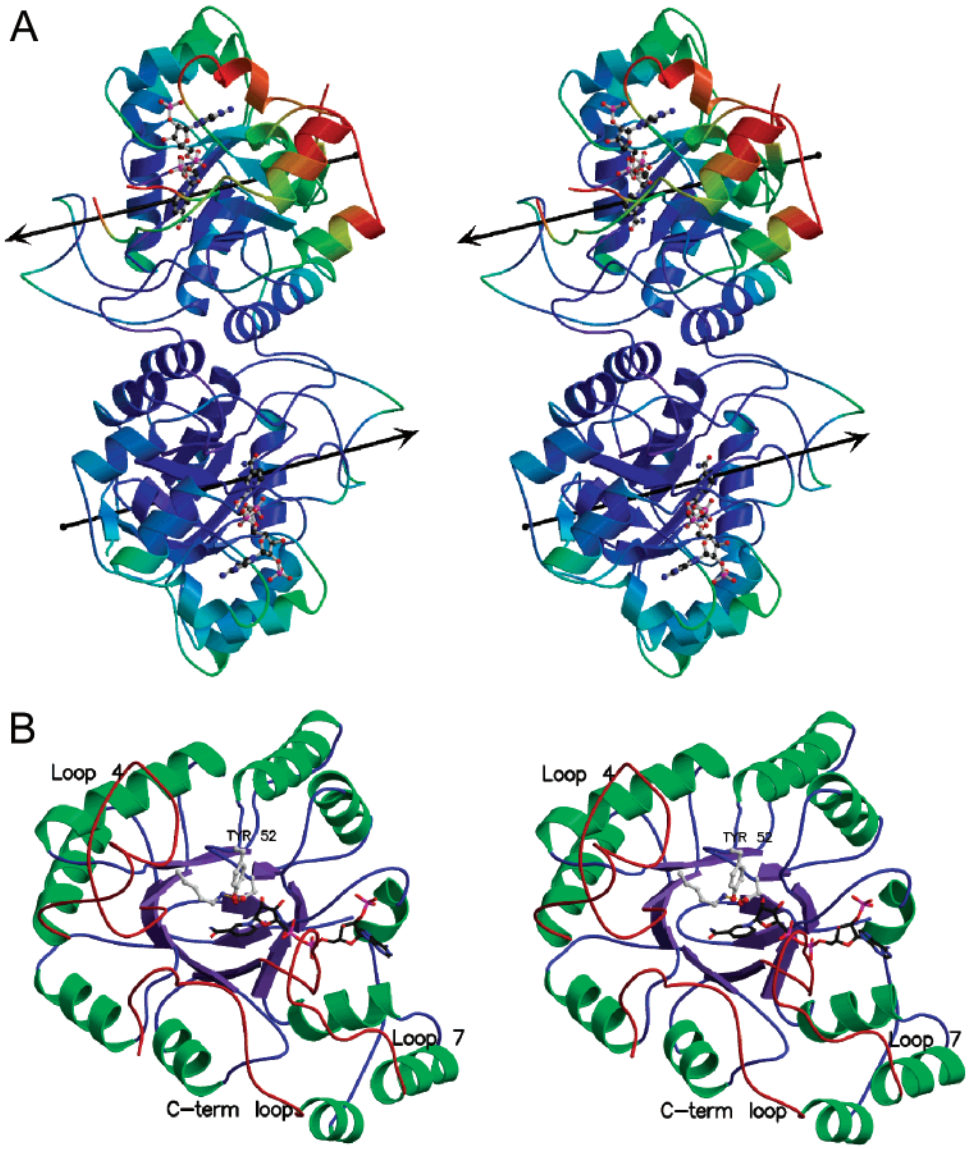

FIGURE 1: (A) Stereoview of the alpha carbon trace of holo XR dimer AB looking down the noncrystallographic 2-fold symmetry axis and colored by temperature factors. Axes of the barrel are depicted by arrows. Colors range from purple for Cα's with temperature factors <15 Å^2 to red for Cα temperature factors >55 Å^2. (B) Stereoview of the alpha carbon trace of holo XR monomer B looking down the β barrel with helices colored green, strands are lavender, and loops are blue. Loops within repeats 4 and 7, and the C-terminal loop are colored crimson. Side chains corresponding to the catalytic residues Asp-47, Lys-81, and Tyr-52 are also shown. The NADPH extends right from the center of the barrel and through the groove formed between repeats 7 and 8.

terminal residues were disordered, as were residues 219−237, which compose loop 7 (the loop between α7 and β7). Apo monomer B has 299 residues modeled. Regions containing residues 1−5 and 224−238 (loop 7) were disordered and not fit. Also found in the apo structures were 407 ordered water molecules which were placed in spherical densities in the F_o-F_c map exceeding 3σ and making at least one plausible hydrogen bond with either protein or other solvent molecules. Examination of a Ramachandran plot as implemented by the program PROCHECK (*20*) indicates that 90.7% of the nonglycine residues fall in the most favored regions, 8.7% in the additionally allowed regions, 0.6% in the generously allowed regions, and none in the disallowed regions.

In the holo asymmetric unit, monomers A and B form dimer 1 while dimer 2 contains monomers C and D. Holo monomer A consists of 317 amino acids. The first three residues as well as residues 227 and 228 were disordered. Holo monomers B, C, and D each consist of 319 amino acids corresponding to residues 4 through 322. There are four NADPH molecules, each associated with one of the four monomers. In addition, the holo model also contains 734 ordered water molecules. There is a notable difference in the overall temperature factors of the holo monomers which vary from 22.89 Å^2 for monomer B to 32.72 Å^2 for monomer D (Table 1). A PROCHECK Ramachandran plot indicates that 89.9% of the nonglycine residues fall in the most favored regions, 9.8% in the additionally allowed regions, 0.3% in the generously allowed regions, and none in the disallowed regions.

In both cases, the secondary structural features are substantially similar. The amino-terminal end of the barrel is capped by a β hairpin turn. The carboxyl end of the internal β barrel contains two loops between the β strand and α helix

[2] Note that amino acid numbering used within includes the initiator methionine as residue number one.

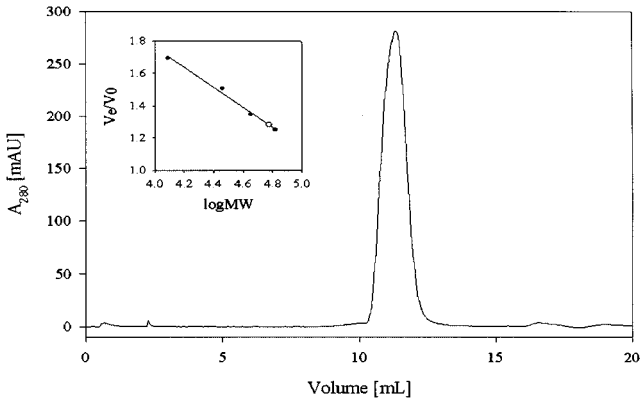

FIGURE 2: Elution profile of XR during gel filtration on a Superdex 75 HR 10/30 sizing column. The inset shows molecular mass determination for XR (open circle) using a calibration obtained with molecular mass standards (full circles). V_e is the elution volume of the sample or the standard, and a value of 7.77 mL for the null volume V_0 of the column was provided by the supplier. The calculated molecular weight for XR is 60 000.

in β/α repeats 4 and 7. Another loop is found at the C-terminus and is observed in other aldo-keto reductase structures. Two extra helices inserted before α7 and after α8 also interrupt the regularity of the structure (Figure 1B). Another common feature present is a groove beginning in the center of the carboxyl terminal end of the β barrel and continuing to the edge between β/α repeats 7 and 8, which corresponds to the NADPH binding site. Important components of this groove are two loops that fold over the cosubstrate, similar to those found in some of the other AKR structures. These two segments interact over the cosubstrate-binding site via two hydrogen bonds and sequester a large region of the NADPH from solvent in the holo structure.

Dimeric Interface. Xylose reductases have been reported to be monomeric in *Saccharomyces cerevisiae* (*21*), and dimeric in *Neurospora crassa* (*22*) and *Pichia stipitis* (*23*). Sizing column experiments have clearly demonstrated that *C. tenuis* XR exists exclusively as a dimer in solution (Figure 2). Dynamic light scattering experiments indicate a monodisperse molecular mass of 69.7 kDa, confirming the dimeric nature of the protein. In all enzymes that have been kinetically characterized, no cooperativity has been observed. Nothing is currently known about the nature of the dimeric interface or whether the formation of this interface is a prerequisite for activity. Examination of the structure suggests that the dimeric interface is probably necessary for structural stability of the protein rather than for catalysis.

The crystallographic asymmetric unit contains one (apo) or two XR dimers (holo), with the dimer(s) being generated by noncrystallographic 2-fold symmetry. The monomers are arranged in a side-to-side fashion with the axes of the barrels approximately 30° of being antiparallel (Figure 1A). A least-squares fit using LSQKAB as implemented in the CCP4 suite of programs (*24*) shows that in all cases, the molecules are very similar. The rms deviations between the alpha carbons of the dimers are 0.50 Å for the holo A−B versus holo C−D and 0.73 and 0.88 Å for the apo versus the holo A−B and holo C−D dimers, respectively. R.m.s. deviations when comparing all atoms are 0.72, 1.10, and 1.23 Å for the same pairs of dimers. The structures of the apo and holo forms of the enzyme are therefore substantially similar. The major structural difference between the two is a conformational change in a loop involved in cosubstrate binding (discussed below). Also noted are minor side chain conformational differences in the cosubstrate binding site.

The intersubunit interaction buries 2328 Å^2 of the total surface area in the apo dimer. Since this is much larger in comparison to the areas mediating crystallographic contacts, it is presumed to represent the structure of the dimer. Virtually identical interactions are observed in the structures of the holo dimers, sequestering 2297 Å^2 from solvent in the A−B dimer and 2335 Å^2 in the C−D dimer. Since the apo and holo structures were determined from two different crystal forms utilizing different lattice contacts, this confirms the relevance of these protein surfaces.

The interface between the monomers is mediated mainly by hydrophilic interactions, but some hydrophobic contacts are present also (Figure 3). Helix α5 is the most important secondary structural element providing these interactions. Other important contributions are made by residues at the end of loop 4 (between β4 and α4), α6 and the end of the carboxyl terminal meander (Table 2).

An analysis of the apo structure indicates that the dimer interface is largely hydrated with 11 buried waters and an additional 12 solvent-accessible waters that mediate hydrogen bonding between the monomers. Only six direct hydrogen bonds involving residues Pro-116, Ile-147, Gly-173, Asp178, Arg181, and Val-322 are observed between the two subunits

Hydrophobic interfaces, which often mediate quaternary interactions (*25*), are limited. Primarily, these consist of Leu-177 from α5 and Lys-202 and Phe-206 from α6 forming a hydrophobic patch that contacts Trp-313, Ile-320, Pro-319, and Val-322 from the C-terminus of the other monomer. A key interaction in this hydrophobic region producing a large number of contacts (Table 2) is the stacking between the aliphatic portion of Lys-202 and the Trp-313 side chain. As a result of the noncrystallographic 2-fold symmetry, both the hydrophobic and the hydrophilic interactions are preserved for the most part on both copies.

The interface observed in the two dimers present in the holo structure demonstrates that the interface is virtually identical when compared to the apo dimer. Side chain conformations and quaternary interactions are nearly indistinguishable. Even water molecules are for the most part conserved, except for two solvent-accessible waters that mediate hydrogen bonds between apo monomers that are not present at the holo dimer interface, and one buried solvent molecule in the apo dimer (that does not mediate hydrogen bonds) that is absent in the holo dimer.

The prototypical $(\beta/\alpha)_8$ barrel is often considered to be the dimeric enzyme triosephosphate isomerase (TIM) (*26*). A comparison of TIM with XR indicates that their modes of dimerization are different both at the gross and detailed levels. TIM dimerizes using residues derived from α2, α3, and loop 3 (the loop preceding α3) and over 1500 Å^2 of surface area are buried per subunit upon dimerization. This corresponds to 16−20% of accessible surface area buried depending on the organism the enzyme is derived from (*25*). In contrast, XR dimerizes by utilizing residues from α4, loop 4, α5, α6, and the C-terminal loop, burying 1164 Å^2 of surface area per subunit corresponding to approximately 9% of the accessible surface area. When the nature of the TIM dimeric interface is examined, the hydrophobic interactions

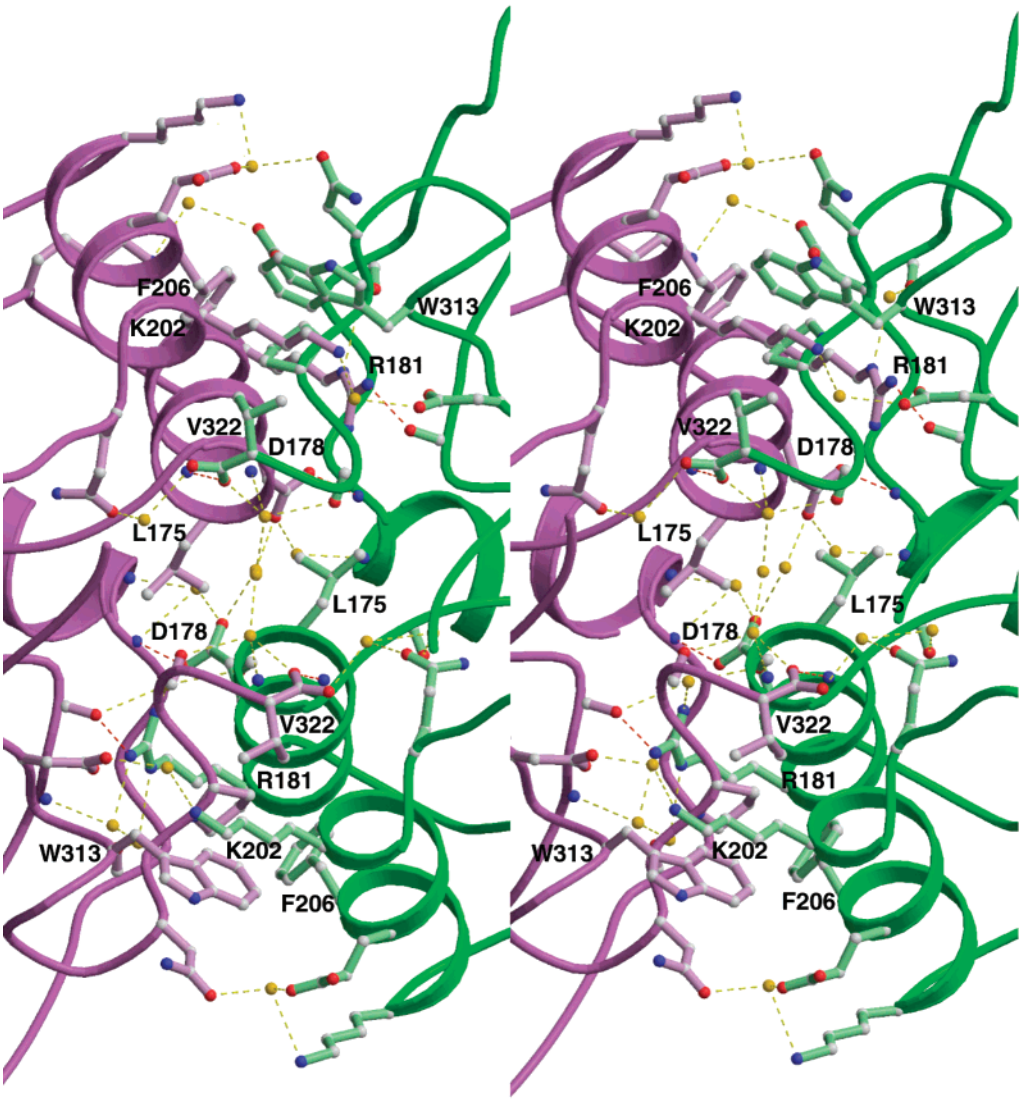

FIGURE 3: A stereoview of the interactions mediating the apo dimer interface. Residues with the most contacts across the interface are labeled. Water molecules and the intricate solvent-mediated hydrogen-bonding network are shown in gold. The six direct hydrogen bonds between the monomers are shown in orange.

Table 2: Contributions of Various Residues to the Dimeric Interface in the Apoenzyme[a]

residue	secondary structural element	monomer A VDW contacts	monomer A hydrogen bonds (direct/indirect)	monomer B VDW contacts	monomer B hydrogen bonds (direct/indirect)
Pro-116	loop 4	0	1/1	0	1/1
Asp-144	loop 4	2	0/0	4	0/1
Pro-146	loop 4	5	0/0	4	0/0
Ile-147	α4	5	1/0	5	1/0
Pro-172	α5	2	0/0	2	0/0
Gly-173	α5	2	1/0	2	1/0
Ala-174	α5	3	0/0	3	0/0
Leu-175	α5	2	0/0	2	0/0
Asp-178	α5	9	1/0	9	1/2
Arg-181	α5	10	1/2	13	1/3
Lys-202	α6	20	0/1	16	0/1
Leu-203	α6	2	0/0	3	0/0
Glu-205	α6	2	0/0	2	0/1
Phe-206	α6	10	0/0	10	0/0
Asp-311	C-terminal loop	0	0/1	0	0/1
Trp-313	C-terminal loop	19	0/0	21	0/0
Asn-317	C-terminal loop	1	0/1	1	0/1
Pro-319	C-terminal loop	10	0/0	10	0/0
Val-322	C-terminal loop	4	1/2	5	1/2

[a] van der Waals interactions are defined as side chain atoms (C_α's included) making < 4 Å contact with any atom on the other subunit. Hydrogen bonds are < 3.2 Å between proton donor and acceptor atoms. Indirect hydrogen bonds are mediated by a single water molecule between donor and acceptor atoms.

Table 3: Conservation of Residues Responsible for Dimerization in *Candida Tenuis* XR among Yeast Monomeric and Dimeric Aldo-Keto Reductases[a]

Enzyme	Loop 4	Helix α5	Helix α6	C-terminal loop
ctenXR (dimer)	^{144}DVPI	^{172}PGALLLDLLR	^{202}KLIEF	^{311}DPWDWDNIPIFV
psXR (dimer)	^{140}DVPI	^{168}PGALLLDLLR	^{198}RLIEF	^{307}DPWDWDKIPIFV
ctropXR (dimer)	^{146}DVPL	^{174}TGALIYDLIR	^{204}KLIEY	^{313}DPWDWDNIPIFV
scGcy1p (monomer)	^{144}NWNF	^{172}SINNLKDLLA	^{204}ELINF	^{299}RVVHPNWSPFEVFK
scXR (monomer)	^{145}HVPI	^{173}QGSLIQDLLR	^{203}HLVEF	^{314}DPWTWLDGKFPTFA

[a] Underlined residues are those making direct intersubunit contacts < 4.0 Å in the *Candida tenuis* XR structure. Residues highlighted in yellow are identical to these. ctenXR: *Candida tenuis* xylose reductase; psXR: *Pichia stipitis* xylose reductase; ctropXR: *Candida tropicalis* xylose reductase; scGcy1p: *Saccharomyces cerevisiae* GCY1 gene product; scXR: *Saccharomyces cerevisiae* xylose reductase.

clearly predominate. Therefore, XRs dimeric interface is not only different in nature when compared to TIM, but the dimeric interaction is much less extensive. Although it remains to be determined whether dimerization is necessary for XR activity, there is no apparent chemical reason the XR monomer could not be catalytically active. In fact, since water molecules extensively mediate the dimeric interface, the monomeric form of the enzyme and the resulting exposed interface could very well be stable in solution.

Oligomerization of other AKRs. Except for the structure of a K^+ channel β subunit (Kvβ2) (*27*), which is a tetramer, all aldo-keto reductase structures determined to date have been monomeric. Aside from the obvious difference in symmetry, the quaternary interactions between subunits are different in XR when compared to Kvβ2. The Kvβ2 protomers associate by allowing the side of the $(\beta/\alpha)_8$ barrel to interact with the amino terminal end of the next barrel in a "head to side" interaction rather than the "side to side" fashion seen in XR.

A comparison of several monomeric and dimeric yeast XR sequences reveals that residues responsible for dimerization are almost completely conserved among those that dimerize and are more divergent in those that are functional as monomers (Table 3). Residues such as Arg-181, Lys-202, Phe-206, Trp-313, and Pro-319, which account for a relatively large number of contacts (Table 2), are either strictly conserved or have very conservative substitutions among the dimers.

In an attempt to identify any other proteins in the superfamily that might function as a homologous dimer, a broader examination of AKR sequence alignments was conducted. This failed to locate any other enzymes that show sequence conservation in these regions. In particular, aflatoxin reductase, an enzyme that may be functional as a dimer (*28*), showed no homology in the residues responsible for dimerization. We therefore conclude that these dimeric AKRs either form their interfaces on other regions of the protein surface or the same surface as XR is used with very divergent amino acids mediating the contacts.

Cosubstrate Binding. The current holo structures show that the NAD(P)H binding pocket is substantially structurally conserved and the cofactor is bound in the same extended anti conformation that is seen in other AKR structures (*15*, *27*, *29*−*31*). Comparing the apo and holo structures, loop 7 (residues 218−242, between β7 and α7) is found to become ordered upon cofactor binding and is held in place by hydrogen bonds between Lys-25 and Ser-224 (Figure 4). There are a large number of specific interactions involved in cosubstrate binding (Figure 5). Initial modeling of the cosubstrate into the binding site of the apo structure indicated that nucleotide binding to XR would require movement of Lys-274 and Ser-275, which occupy the region where the pyrophosphate should be located. Rotation of the side chain of Trp-24 to a position roughly perpendicular to the pyrophosphate is also required to remove any steric conflict (Figure 4). The holoenzyme structure confirmed the perturbations predicted from docking an NADPH molecule into the apoenzyme structure. In fact, the alpha carbon of Lys-274 shifts 3.5 Å upon nucleotide binding. The side chain of Tyr-217 stacks with the nicotinamide ring providing numerous contacts. The side chains of Ser-169, Asn-170, and Gln-191 orient the 4-*pro-R* face of the nicotinamide ring toward the opening of the cavity by contributing hydrogen bonds to the amide group. These interactions coupled with the stacking atop Tyr-217 are consistent with the observation that the enzyme promotes a transfer that is stereospecific for the 4-*pro-R* hydride. This agrees with biochemical studies demonstrating that XR is a reductase specific for the hydride residing on the A side of the nicotinamide (*32*). Conservation of the serine and asparagine residues throughout the majority of the AKR family strongly suggests that the stereospecificity of hydride transfer is also conserved among these members. The structure of the more divergent Kvβ2 reveals that the interaction with a serine is conserved, but the asparagine is replaced by an arginine. Greater sequence divergence in the hydride specificity region and the amino acids responsible for nicotinamide orientation is observed in AKR families 6, 7, 8, 9, 10, 11, and 12, suggesting that transfer may not always be specific for the 4-*pro-R* hydride. This is indeed the case with a human succinic semialdehyde reductase (also know as aflatoxin reductase), a member of family 7, which has been characterized as being specific for the 4-*pro-S* hydride (*33*).

While most of the residues that contribute hydrogen bonds and salt links to the cofactor are conserved in hAR and XR, there are two notable exceptions: Ser-214 in hAR, which contributes main chain amide and side chain hydrogen bonds to the pyrophosphate is replaced by Pro-222 in XR, and Thr-265 in hAR, which contributes a side chain hydrogen bond to the 2′ phosphate of the adenosine is replaced by Leu-277 in XR. These two substitutions as well as a lack of the salt-linked "safety belt" seen in hAR probably contribute to the lower affinity for NADPH in XR versus hAR. This hydrophobic residue at position 277 is unique in any of the AKR

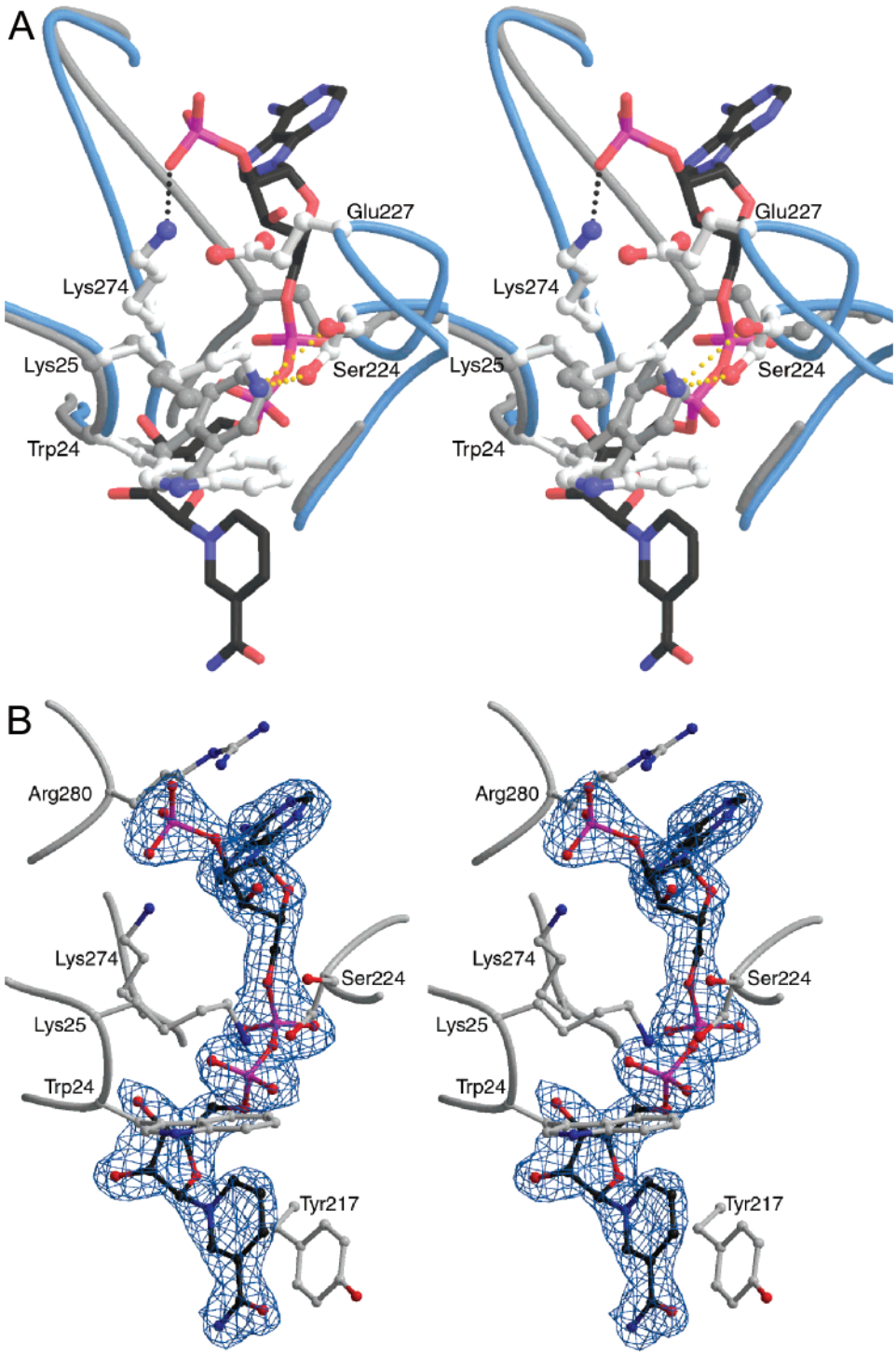

FIGURE 4: (A) Stereoview of the main chain ordering and conformational changes involved in cosubstrate binding. Apo monomer A (grey) has been superpositioned on holo monomer B (blue with side chains colored by atom type). Loop fastening hydrogen bonds between Lys-25 and Ser-224 are represented as dashed yellow lines. Lys-274 is shown interacting with the NADPH 2′ adenosine phosphate. Loop 7, which contains Ser-224, is disordered in the apo structure. (B) Representative $2F_o-F_c$ electron density surrounding the cosubstrate in an orientation similar to panel A. Density is a simulated annealing omit map contoured at 2σ.

structures so far determined and may allow the dual cofactor specificity that has been observed in XR (*34*).

Catalysis and Substrate Binding Site. The active site cavity of the apo XR enzyme is approximately 8.5 × 11 Å wide and 14 Å deep. Positioned near the bottom of the active site pocket are side chains from Tyr-52, Asp-47, and Lys-81 that are poised to form a catalytic triad activating the tyrosine to function as the general acid (*15*). The proton on the phenolic oxygen will have a depressed pK_a as a result of interactions with Nζ on Lys-81. Lys-81 in turn engages in a salt link with the Asp-47 side chain carboxylate. These roles have been confirmed by assaying relevant mutants in hAR (*35*, *36*). His-114 is believed to serve in orienting the substrate carbonyl, a function analogous to His-110 in human aldose reductase (*36*).

The apparent affinity (K_m = 87 mM) for the xylose substrate is explained structurally by examining the nature of the substrate-binding pocket. It is predominantly hydro-

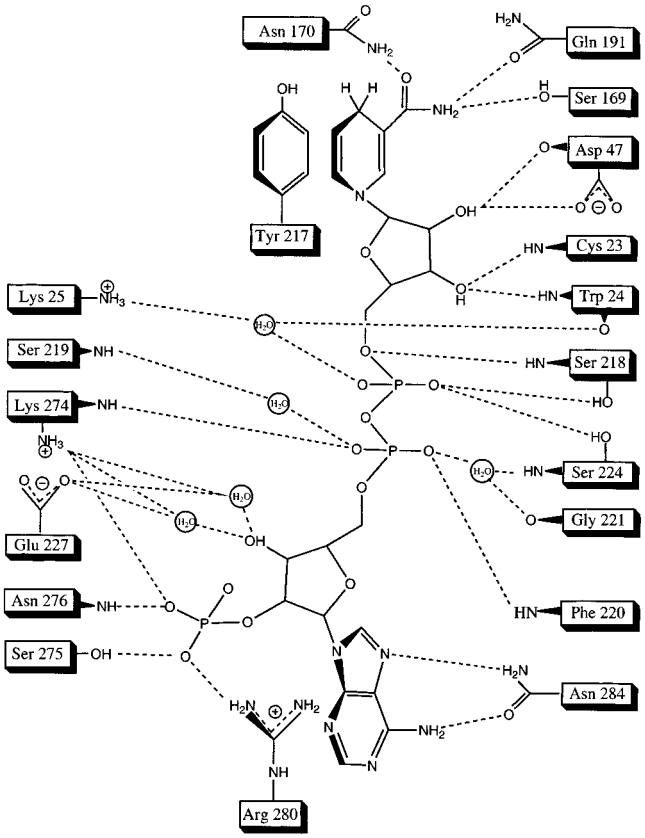

FIGURE 5: A schematic representation of interactions mediating NADPH binding to XR taken from holo monomer B.

phobic and lined with Trp-24, Trp-83, Phe-115, Pro-129, Phe-132, Pro-312, and Trp-315 in the apo structure. When comparing the apo and holo forms of the enzyme Phe-225, Met-228, and Gln-230 are additionally seen in the substrate-binding site as a result of loop 7 (residues 218−242) becoming ordered upon binding of the cosubstrate (Figure 6). Polar residues able to engage in hydrogen bonding, a necessary component for a high-affinity xylose binding site, are slightly more numerous than in other AKRs but too few to compose an high affinity xylose binding site. Two residues likely to be accessible to a carbonyl oxygen in the substrate include Tyr-52 of the catalytic triad and His-114. Other polar residues possibly involved in binding are Asp-51, Gln-230, and Asn-310, which are interspersed among the aromatic and aliphatic side chains (Figure 6). Asn-310 in XR is structurally homologous to Cys-298 in human aldose reductase. This cysteine has been identified as the target of inactivation in hAR by oxidizing physiological disulfides such as GSSG and cystine (*37*). This inactivation is not observed in XR (*34*) and is very likely to be due to the replacement of this solvent accessible cysteine. Gln-230 resides at the lip of the cavity approximately 10 Å from the nicotinamide C4 and may be too far removed to contribute substantially to substrate binding, particularly for smaller polar substrates. Asp-51 contributes most to the relative increase in binding site polarity since the structurally analogous residue is an Ile, Leu, or Ala in all previous AKR structures. Despite the hydrophobic character of the substrate-binding site, seven and five ordered water molecules are observed in each of the apo monomer binding sites and six, five, five, and four are seen in the active sites of the holo enzyme. There is additional density near the Asp-51 carboxylate greater than 5.5 σ on a F_o-F_c map that was not accounted for in the model for holo monomers B, C, and D and apo monomer B.

To gain an understanding of how XR might bind its aldehydic substrate, modeling experiments were conducted to place open chain xylose into the active site. Ordered water molecules found in the holo structure were used as a guide to manually dock the substrate so the C1 was within hydride transfer distance above the nicotinamide C4 and the carbonyl oxygen hydrogen-bonded to the phenolic oxygen of Tyr-52 and Nϵ2 of His-114. This is equivalent to the calculated carbonyl position for the binding of D-xylose to hAR (*38*). In this orientation, the aldehyde proton points toward Trp-24, and the *re* face of the carbonyl is positioned toward and parallel with the nicotinamide ring consistent with the predicted stereochemistry of the reaction (*6*). The side chains of residues Asp-51 and Asn-310 are subsequently available to bind the xylose C2 hydroxyl. Specific interactions between XR and the C2(R) hydroxyl group are important to transition state stabilization (*6*). Asp-51 contributes the most to the relative polarity of the binding site, but attempts to hydrogen

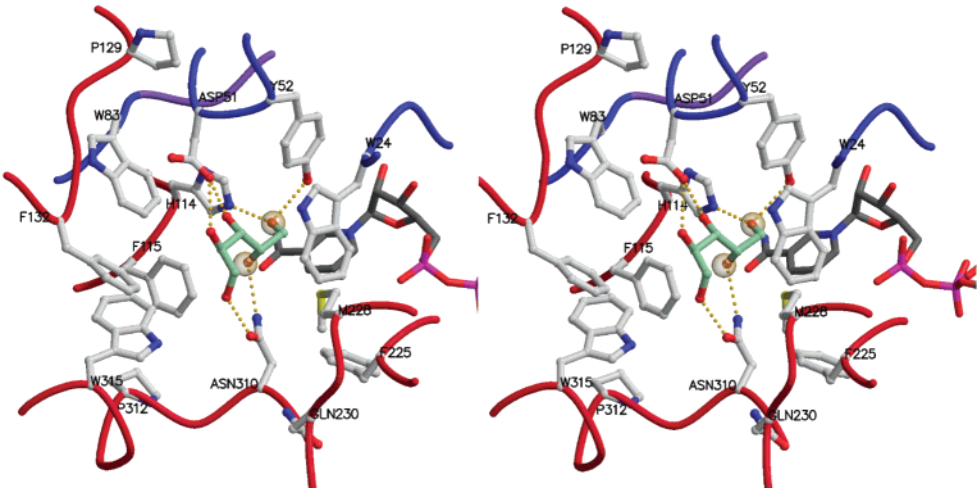

FIGURE 6: Stereoview of the XR active site with the xylose substrate (green carbons, red oxygens) modeled. The polar amino acids not present in other AKRs are labeled in three-letter code. Secondary structure is colored as in Figure 1B, and the NADPH cofactor in gray extends to the right. Two crystallographically observed water molecules are represented as transparent yellow spheres and possible hydrogen bonds between xylose and the enzyme are shown as dashed yellow lines.

bond this hydroxyl to Asp-51 resulted in steric conflict between Trp-24 and the remainder of the xylose chain. Alternatively, if the C2 hydroxyl is placed within hydrogen bond distance of the side chain nitrogen of Asn-310 no steric collisions are observed, and several potential hydrogen bonds can be made between side chains of Trp-24, Asp-51, Asn-310 and hydroxyls on xylose carbons 3, 4, and 5. We therefore predict that the xylose C2 hydroxyl forms a hydrogen bond to Nδ of Asn-310. This is in agreement with data suggesting that the C2 hydroxyl acts primarily as an acceptor of hydrogen bonds (*6*). In a series of experiments designed to elucidate how XR uses noncovalent interactions to bind its substrate, Neuhauser et al. measured activity with stereoisomers of polyhydroxylated aldehydes and deoxy and deoxyfluoro analogues of galactose. They concluded (i) that hydrogen bonding between XR and the C2(R) hydroxyl is crucial to transition state binding and can be partially replaced by fluorine suggesting that the C2 hydroxyl acts primarily as a hydrogen bond acceptor, and (ii) substitution of hydroxyls by hydrogens at other carbon position had no effect on catalytic efficiency, but correct stereochemistry at carbon positions 3−5 avoids nonfavorable interactions. A binding conformation consistent with these findings is shown in Figure 6 with the carbonyl oxygen and C2 hydroxyl superimposing upon ordered water molecules found in the holo structures.

Although we are unaware of any experiments that have looked at inhibition of XR by zopolrestat, a well-studied and potent inhibitor of most other AKRs, modeling attempts placing zopolrestat into the enzyme substrate-binding site were not successful. Binding in the same position that is seen in hAR and another AKR, FR-1 (murine fibroblast growth factor induced protein-1) would bring the benzyl portion of the phthalazinone ring within 2.2 Å of Asp-51 and would require insertion of the tri-fluoro methyl group into a tight hydrophobic pocket made up of residues Pro-116, Ile-117, Ile-318, Pro-312, and Ile-320 near the C-terminus. These residues are held in place by a hydrophobic dimer interaction. It should be noted that modeling studies of zopolrestat and other AKR inhibitors have been unsuccessful in the past using holoenzyme structures due to the plasticity of the active site. Crystal structures of the hAR holoenzyme (*15*) and hAR holoenzyme-zopolrestat ternary complex (*39*) demonstrate large conformational changes upon zopolrestat binding. Such unpredictable changes are possible and perhaps even likely upon an inhibitor binding to XR.

Conformational Flexibility in Cosubstrate Binding. In the forward (reducing) direction, some AKRs such as human aldose reductase appear to have a rate-limiting step that is associated with either the dissociation of the consumed cosubstrate or a conformational change linked to the cosubstrate dissociation (*40*), while others such as 3α-hydroxysteroid dehydrogenase (3αHSD) exhibit other rate-limiting steps. Kinetic results for XR demonstrate that the dissociation of $NAD(P)^+$ is faster than that seen in human aldose reductase but slower than in 3αHSD, so that this step still appears to be rate-limiting (*41*). This can be explained by the composition of loop 7, one of the loops responsible for enfolding the cosubstrate. In hAR, residues 213−217 at the beginning of loop 7, fold over the pyrophosphate and are held tightly in place by a bidentate salt link between Asp-216 on the loop and Lys-21 and Lys-262 on β1 and β8, respectively. Examination of aligned sequences shows that XR has a glutamate present at position 227 homologous to Asp-216 in hAR. Upon the basis of sequence information alone, it is plausible that Lys-25 and Lys-274 would fulfill a role analogous to Lys-21 and Lys-262 in the human form. This is not what is observed in the holo XR structure, however. The alpha carbons of Lys-274 and Glu-227 are situated too far apart (> 13 Å) for their side chains to interact directly. In lieu of a direct interaction, there are two solvent mediated contacts between Lys-274 and Glu-227 creating a safety belt that is more weakly fastened. The Nζ of Lys-274 also interacts with the 2′ adenosine phosphate group. Lys-25 is close enough to interact with Glu-227 but is oriented toward Trp-24 instead, and the loop is alternatively fastened by hydrogen bonds between the Nζ of Lys-25 and the side chain hydroxyl and main chain carbonyl of Ser-224 (Figure 4). This weaker interaction relative to that found in hAR and some other AKRs may explain the experimentally observed 200-fold increase in the ctXR rate constant describing the conformational isomerization preceding cosubstrate release (*42*). AKRs such as 3αHSD have rate-limiting steps that are not determined by dissociation of the cosubstrate (*43*). The enzymes in this class are either missing this loop, or, in the case of 3αHSD, this loop is fastened via weaker, nonhomologous interactions (*44*).

There has also been plasticity observed in the substrate binding site of AKRs. Structural studies of the 2,5-diketo-D-gluconic acid reductase apoenzyme show that cosubstrate binding causes large conformational rearrangements in the active site. The NADPH is therefore not only chemically necessary for catalysis but also for proper placement of side chains necessary for catalysis and substrate binding (*45*). A similar isomerization may also occur in Gcy1p, an AKR from *Saccharomyces cerevisiae* (*46*). XR shows no such large scale movement indicating that this is not a general phenomenon in AKRs.

The structure of XR demonstrates a new quaternary organization for aldo-keto reductases which is likely to be conserved among the xylose and mannose reductases but not among more distantly related dimeric families. Although its catalytic mechanism is identical to that found in other AKRs, the substrate binding site is more hydrophilic, allowing for specific hydrogen bond formation with xylose. The structure also defines how the enzyme recognizes its NADPH cosubstrate and suggests how the specificity might be altered to produce an exclusively NADH-dependent activity that is more amenable to in vivo metabolism of xylose.

REFERENCES

1. Jez, J. M., Bennett, M. J., Schlegel, B. P., Lewis, M., and Penning, T. M. (1997) *Biochem. J. 326*, 625−636.
2. Lee, H. (1998) *Yeast 14*, 977−984.
3. Aspinall, G. O. (1980) in *The Biochemistry of Plants* (Preiss, J., Stumpf, P. K., and Conn, E. E., Eds.) pp 473−499, Academic Press, New York.
4. Hahn-Hagerdal, B., Linden, T., Senac, T., and Skoog, K. (1991) *Appl. Biochem. Biotechnol. 28−29*, 131−144.
5. Nidetzky, B., Neuhauser, W., Haltrich, H., and Kulbe, K. D. (1995) *Biotechnol. Bioeng. 52*, 387−396.
6. Neuhauser, W., Haltrich, H., Kulbe, K. D., and Nidetzky, B. (1998) *Biochemistry 37*, 1116−1123.
7. Grimshaw, C. E. (1992) *Biochemistry 31*, 10139−10145.
8. Bedford, J. J., Bagnasco, S. M., Kador, P. F., Harris, H. W., and Burg, M. B. (1987) *J. Biol. Chem. 262*, 14255−14259.

9. Kador, P. F., Robinson, W. G., and Kinoshita, J. H. (1984) *Annu. Rev. Pharmacol. Toxicol. 25*, 691−714.
10. Haecker, B., Habenicht, A., Kiess, M., and Mattes, R. (1999) *Biol. Chem. 380*, 1395−1403.
11. Mayr, P., Brueggler, K., Kulbe, K. D., and Nidetzky, B. (2000) *J. Chromatogr. B 737*, 195−202.
12. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307−326.
13. Matthews, B. W. (1968) *J. Mol. Biol. 33*, 491−497.
14. Gehlhaar, D. K., and Fogel, D. B. (1999) *Acta Crystallogr. D55*, 484−491.
15. Wilson, D. K., Bohren, K. M., Gabbay, K. H., and Quiocho, F. A. (1992) *Science 257*, 81−84.
16. Sack, J. S. (1988) *J. Mol. Graphics 6*, 224.
17. Jones, T. A., Zou, J.-Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
18. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D54*, 905−921.
19. Branden, C.-I. (1991) *Curr. Opin. Stuct. Biol. 1*, 978−983.
20. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
21. Kuhn, A., van Zyl, C., van Tonder, A., and Prior, B. A. (1995) *Appl. Environ. Microbiol. 61*, 1580−1585.
22. Rawat, U. B., and Rao, M. B. (1996) *Biochim. Biophys. Acta 1293*.
23. Rizzi, M., Erlemann, P., Bui-Thanh, N. A., and Dellweg, H. (1988) *Appl. Microbiol. Biotechnol. 29*, 148−154.
24. Collaborative Computational Crystallography Project, N. (1994) *Acta Crystallogr. D50*, 760−763.
25. Delboni, L. F., Mande, S. C., Rentier-Delrue, F., Mainfroid, V., Turley, S., Vellieux, F. M. D., Martial, J. A., and Hol, W. G. J. (1995) *Protein Sci. 4*, 2594−2604.
26. Waley, S. G. (1973) *Biochem. J. 135*, 165−172.
27. Gulbis, J. M., Mann, S., and MacKinnon, R. (1999) *Cell 97*, 943−952.
28. Kelly, V. P., Ireland, L. S., Ellis, E. M., and Hayes, J. D. (2000) *Biochem. J. 348*, 389−400.
29. Wilson, D. K., Nakano, T., Petrash, J. M., and Quiocho, F. A. (1995) *Biochemistry 34*, 14323−14330.
30. Ye, Q., Hyndman, D., Li, X., Flynn, T. G., and Jia, Z. (2000) *Proteins: Struct., Funct. Genet. 38*, 41−48.
31. Khurana, S., Powers, D. B., Anderson, S., and Blaber, M. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 6768−6773.
32. Nidetzky, B., Mayr, P., Neuhauser, W., and Puchberger, M. (2001) *Chem.-Biol. Interact. 130*−*132*, 583−595.
33. Hoffman, P. L., Wermuth, B., and von Wartburg, J.-P. (1980) *J. Neurochem. 35*, 354−366.
34. Neuhauser, W., Haltrich, D., Kulbe, K. D., and Nidetzky, B. (1997) *Biochem. J. 326*, 683−692.
35. Tarle, I., Borhani, D. W., Wilson, D. K., Quiocho, F. A., and Petrash, J. M. (1993) *J. Biol. Chem. 268*, 25687−25693.
36. 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.
37. Cappiello, M., Voltarelli, M., Cecconi, I., Vilardo, P. G., Massimo, D. M., Marini, I., Del Corso, A., Wilson, D. K., Quiocho, F. A., Petrash, J. M., and Mura, U. (1996) *J. Biol. Chem. 271*, 33539−33544.
38. De Winter, H. L., and von Itzstein, M. (1995) *Biochemistry 34*, 8299−8308.
39. Wilson, D. K., Tarle, I., Petrash, J. M., and Quiocho, F. A. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 9847−9851.
40. Grimshaw, C. E., Bohren, K. M., Lai, C.-J., and Gabbay, K. H. (1995) *Biochemistry 34*, 14356−14365.
41. Nidetzky, B., Mayr, P., Hadwiger, P., and Stuetz, A. E. (1999) *Biochem. J. 344*, 101−107.
42. Nidetzky, B., Klimacek, M., and Mayr, P. (2001) *Biochemistry 40*, 10371−10381.
43. Ratman, K., Ma, H., and Penning, T. M. (1999) *Biochemistry 38*, 7856−7864.
44. Bennett, M. J., Schlegel, B. P., Jez, J. M., Penning, T. M., and Lewis, M. (1996) *Biochemistry 35*, 10702−10711.
45. Sanli, G., and Blaber, M. (2001) *J. Mol. Biol. 309*, 1209−1218.
46. Hur, E., and Wilson, D. K. (2001) *Chem.-Biol. Interact. 130*, 527−536.

BI025786N