1 Structural characterization of an aldo-keto reductase (AKR2E5) from the
2 silkworm *Bombyx mori*

3
4 Kohji Yamamoto ^{a,*} Akifumi Higashiura ^b, Mamoru Suzuki ^b, Takahiro Shiotsuki ^c,
5 Ryohei Sugahara ^c, Takeshi Fujii ^{d,e}, Atsushi Nakagawa ^b

6
7 ^a *Kyushu University Graduate School, 6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-*
8 *8581, Japan*

9 ^b *Institute for Protein Research, Osaka University, Suita 565-0871, Japan*

10 ^c *National Institute of Agrobiological Sciences, Tsukuba, Ibaraki 305-8634, Japan*

11 ^d *Graduate School of Agricultural and Life Sciences, The University of Tokyo, Tokyo*
12 *113-8657, Japan*

13 ^e *Department of Science of Biological Production, Graduate School of Agriculture,*
14 *Tokyo University of Agriculture and Technology, Fuchu, Tokyo 183-8509, Japan*

15
16 * Corresponding author. Fax: +81 92 6241011

17 E-mail address: yamamok@agr.kyushu-u.ac.jp

18
19 Abbreviations used: AKR, aldo-keto reductase; DLG, DL-glyceraldehyde; PCR,
20 polymerase chain reaction; RT-PCR, reverse transcriptase PCR; SDS-PAGE, sodium
21 dodecyl sulfate polyacrylamide gel electrophoresis; TIM, triose phosphate isomerase

22
23 *Key words:* protein crystal structure; enzyme specificity; NADPH; Lepidoptera

1

1 A B S T R A C T

2 We report a new member of the aldo-keto reductase (AKR) superfamily in the silkworm
3 *Bombyx mori*. Based on its amino acid sequence, the new enzyme belongs to the AKR2
4 family and was previously assigned the systematic name AKR2E5. In the present study,
5 recombinant AKR2E5 was expressed, purified to homogeneity, and characterized. The
6 X-ray crystal structures were determined at 2.2 Å for the apoenzyme and at 2.3 Å
7 resolution for the NADPH-AKR2E5 complex. Our results demonstrate that AKR2E5 is
8 a 40-kDa monomer and includes the TIM- or (β/α)₈-barrel typical for other AKRs. We
9 found that AKR2E5 uses NADPH as a cosubstrate to reduce carbonyl compounds such
10 as DL-glyceraldehyde, xylose, 3-hydroxy benzaldehyde, 17 α -hydroxy progesterone, 11-
11 hexadecenal, and bombykal. No NADH-dependent activity was detected. Site-directed
12 mutagenesis of AKR2E5 indicates that amino acid residues Asp70, Tyr75, Lys104, and
13 His137 contribute to catalytic activity, which is consistent with the data on other AKRs.
14 To the best of our knowledge, AKR2E5 is only the second AKR characterized in
15 silkworm. Our data should contribute to further understanding of the functional activity
16 of insect AKRs.

1. Introduction

The aldo-keto reductase (AKR) superfamily comprises ubiquitously expressed enzymes responsible for the reduction of aldehydes, ketones, steroid hormones, and monosaccharides [1], mainly via NADPH-dependent reversible reactions. Multiple AKR families and subfamilies have been defined based on the differences in amino acid sequences [2], which influence substrate specificity. For example, the AKR family 1 (AKR1) comprises hydroxysteroid dehydrogenases, whereas the AKR2 family contains aldose reductases. Other AKR families are involved in prostaglandin synthesis, xenobiotic detoxification, reduction of phospholipid aldehydes, metabolic assimilation of carbohydrates, and other secondary metabolic pathways [2]. In all characterized members, AKR structures contain the triose phosphate isomerase (α/β)₈-barrel (TIM barrel) motif with a cylindrical core of eight parallel β -stands surrounded by eight α -helices running antiparallel to the strands. AKR binding sites are formed when the cofactor NADPH binds to the pocket generated by the C-terminal face of the TIM barrel. Residues involved in cofactor bindings, such as the catalytic tetrad Asp, Tyr, Lys, and His, are highly conserved [3].

In our previous study, we identified mRNA coding for a hitherto unknown AKR in the silkworm (*Bombyx mori*), a model Lepidoptera [4]. Here, we biochemically characterized this enzyme designated AKR2E5. Our kinetic studies demonstrated that AKR2E5 recognized aldehydes, steroid hormones, sugars, and insect pheromones. We also determined AKR2E5 X-ray crystal structure and investigated structure-function relationships in enzyme catalysis using the recombinant AKR2E5 overexpressed in *Escherichia coli*. Our results suggest that conserved amino acid residues in the NADPH-binding region of AKR2E5 are critical for catalysis.

2. Materials and methods

2.1. Insects and tissue dissection

Silkworm larvae were reared on mulberry leaves in the Institute of Genetic Resources at Kyushu University Graduate School (Fukuoka, Japan). Day-3 fifth-instar larvae were dissected on ice, and the resulting tissues were stored at -80°C until use. Total RNA was extracted from the dissected tissues using an RNeasy Plus Mini Kit (Qiagen, Valencia, CA, USA), following the manufacturer's protocol, and then used for RT-PCR.

2.2. Cloning and sequencing of AKR2E5-encoding cDNA (*akr2e5*)

Total RNA was analyzed using RT-PCR. First-strand cDNA was synthesized using SuperScript II Reverse Transcriptase (Invitrogen, Carlsbad, CA, USA) and an oligo-dT primer and used as a template for PCR with the following oligonucleotide primers: 5'-AGAACACATATGAAATAAGGTGAATGTACCC-3' (sense) and 5'-AAGGATCCTTAATCGATGTCACCGTTGGCA-3' (antisense). Primers were designed based on the sequence obtained from the SilkBase EST database [5]; the underlined and double-underlined regions are *Nde*I and *Bam*HI restriction enzyme sites, respectively, used to insert the PCR product into the plasmid. The thermocycling conditions were as follows: 94°C for 2 min, followed by 35 cycles at 94°C for 1 min, 56°C for 1 min, and 72°C for 2 min, with a final extension step at 72°C for 10 min. AKR2E5 cDNA

(*akr2e5*) was ligated into the pGEM-T Easy Vector (Promega) and used to transform *E. coli* TOP10. The clone was sequenced using the GenomeLab DTCS Quick Start Kit (Beckman) and an automated DNA Sequencer (CEQ 8000; Beckman). We used GENETYX-MAC software (ver. 14.0.12) to obtain the complete *akr2e5* gene sequence and to deduce the amino acid sequence.

Sequence identity alignments were performed with ClustalW (ver. 1.83); the gap creation penalty and gap extension were set to 10 and 0.2, respectively.

2.3. Overexpression and purification of the recombinant protein

The pGEM-T Easy Vector carrying *akr2e5* was digested with *Nde*I and *Bam*HI, subcloned into the expression vector pET-11b (Novagen), and transformed into competent *E. coli* Rosetta-gami™ 2 (DE3) cells (Novagen). Cells were grown at 37°C in Luria-Bertani (LB) medium containing 100 µg/mL ampicillin until the optical density (OD₆₀₀) reached 0.7, and induced with 1 mM isopropyl 1-thio-β-D-galactoside (PMSF) for the production of recombinant protein. After further incubation for 3 h, cells were centrifuged, resuspended in 20 mM Tris-HCl containing 0.5 M NaCl (pH 8.0), 4 mg/mL lysozyme, and 1 mM PMSF, and disrupted with sonication.

Unless otherwise noted, all subsequent operations were conducted at 4°C. The disrupted cells were centrifuged at 10,000 × *g* for 15 min, and the resulting supernatant was subjected to ammonium sulfate fractionation (30% to 70% saturation). The precipitate was resuspended in 10 mM Tris-HCl buffer (pH 8.0), dialyzed against the same buffer, and subjected to anion-exchange chromatography on DEAE-Sepharose (GE Healthcare); the bound protein was eluted by a linear gradient (0–0.3 M) of NaCl.

Enzyme-containing fractions were pooled, concentrated using a centrifugal filter (Millipore), applied to a Superdex 200 column (GE Healthcare), and equilibrated in Tris-HCl containing 0.2 M NaCl. SDS-PAGE was performed in a 15% polyacrylamide slab gel containing 0.1% SDS, as previously described [6]. Protein bands were visualized by staining with Coomassie Brilliant Blue R250.

2.4. Crystallization, data collection, and structure determination

Purified AKR2E5 protein was concentrated to approximately 10 mg/mL in 10 mM HEPES (pH 7.4). Crystallization was performed using a sitting-drop vapor diffusion method at 20°C. Each drop was formed by mixing equal volumes (1 μ L) of protein and reservoir solutions. Crystals used for X-ray analysis were grown at 20°C for 1 month in 0.1 M Tris-HCl (pH 8.5) containing 4 M ammonium acetate. X-ray diffraction data sets were collected with beamline BL44XU at SPring-8 (Japan Synchrotron Radiation Research Institute, Hyogo, Japan) under cryogenic conditions. Specifically, crystals were harvested into a cryoprotectant solution composed of 77% (v/v) reservoir solution and 23% (v/v) ethylene glycol. The diffraction data were integrated using the XDS package [7] and scaled using the SCALA software [8]. Before analysis, crystals were soaked in the same reservoir solution containing 0.25 mM NADPH for 20 s.

AKR2E5 structure was determined by a molecular replacement method using the Phaser crystallographic software [9]. A previously reported structure of *B. mori* aldo-keto reductase (AKR2E4, PDB ID: 3WCZ) was used as a search model [10]. AKR2E5 structure was refined by the Phenix software [11], and manual model refitting to the

electron density map was performed using the Coot tool [12]. The stereochemical quality of the final model was assessed using MolProbity [13]. The NADPH-binding site of AKR2E5 was examined by the LIGPLOT program [14]. All figures were prepared using PyMOL (<https://www.pymol.org/citing>). The atomic coordinates and associated structure factors of AKR2E5 were deposited in the PDB (ID: 5AZ0 for apo-AKR2E5 and 5AZ1 for the NADPH-AKR2E5 complex).

2.5. Measurement of enzyme activity

AKR activity was measured in the reaction mixture containing 0.25 mM NADPH or NADH, 50 mM potassium phosphate (pH 7.0), 10 mM KCl, and 0.5 mM EDTA, using a Shimadzu UV-2400PC spectrophotometer (Kyoto, Japan). The decrease in absorbance at 340 nm was monitored at 25°C and converted into NADH or NADPH concentrations ($\epsilon_{340} = 6220 \text{ M}^{-1} \text{ cm}^{-1}$ or $\epsilon_{340} = 6200 \text{ M}^{-1} \text{ cm}^{-1}$, respectively, where ϵ is the molar absorption coefficient). Kinetic parameters were obtained from the Michaelis-Menten equation using the KaleidaGraph software (Synergy Software).

3. Results and discussion

3.1. Identification and sequencing of AKR2E5

In the present study, we identified and characterized AKR2E5, a novel AKR in the silkworm. This study follows our previous research on AKR2E4 [15] and brings the number of fully characterized AKRs in the silkworm to two, increasing our

1 understanding of the physiological role of these enzymes in silkworms.

2 AKR2E5 cDNA was cloned from total RNA of the fat body using RT-PCR. The
3 resulting nucleotide sequence of *akr2e5* has been deposited in GenBank (accession
4 number LC012974). The *akr2e5* gene contains a coding sequence of 1,062 bp that
5 encodes 353 amino acid residues. The theoretical molecular mass and pI (isoelectric
6 point) of AKR2E5 are 39,800 kDa and 5.06, respectively. The systematic name of
7 AKR2E5 was given based on its sequence similarity with the members of the AKR2
8 subfamily E (<http://www.med.upenn.edu/akr/>). A BLAST search revealed that the
9 deduced AKR2E5 amino acid sequence exhibited 69%, 56%, and 38% identity to AKR
10 homologs from African swallowtail, *Papilio dardanus* (UniProt KB: B8ZX07), noctuid
11 moth, *Agrotis ipsilon* (U5KCB8), and cotton bollworm, *Helicoverpa armigera*
12 (*HarmAKR*, F5BYH7), respectively (Supplementary Fig. 1). Previously, we have
13 identified one isozyme from *B. mori* (AKR2E4) [15]. The amino acid sequences of
14 AKR2E5 and AKR2E4 are 38% identical and 62% similar.

16 3.2. X-ray crystal structure of AKR2E5

18 The apoenzyme was crystallized in the space group $P6_122$ with the following unit
19 cell dimensions: $a = 112.12 \text{ \AA}$, $b = 112.12 \text{ \AA}$, $c = 147.26 \text{ \AA}$. The X-ray crystal structure
20 was determined by molecular replacement. The final R_{work} and R_{free} (indicating the
21 quality of the atomic model obtained from crystallographic data) were 15.7% and
22 20.4%, respectively (Table 1), at the resolution of $2.2 \text{ \AA}$. Root-mean-square deviations
23 from the ideal bond lengths and ideal bond angles were $0.021 \text{ \AA}$ and 1.909° ,

1 respectively. The Ramachandran plot indicated that 96.9% of the main-chain dihedral
2 angles were in the preferred regions, with 3.1% in the allowed regions.

3 The NADPH-enzyme complex had the same space group as the apoenzyme.
4 Crystallographic parameters and statistical analysis of the NADPH-enzyme complex are
5 summarized in Table 1. The results of Superpose [16] indicated no significant structural
6 differences between apo-AKR2E5 and NADPH-AKR2E5, as evidenced by the
7 observation that the root-mean-square deviation (C^α) ($0.167 \text{ \AA}/344$ residues) was much
8 lower than the expected coordinate errors. The data in Table 1 suggest that the structure
9 contained ethylene glycol, acetate, and calcium, which could have been acquired from
10 the cryoprotectant, crystallization medium, or *E. coli* culture medium.

11 AKR2E5 structure included eight regular barrel α -helices ($\alpha 1$ – $\alpha 8$) and two
12 auxiliary helices found in other AKRs such as H1 and H2 (Fig. 1). Additionally, the
13 final model contained eight β -strands ($\beta 1$ – $\beta 6$) and two other strands designated as B1 and
14 B2 (Fig. 1). The final enzyme model revealed that the apo-AKR2E5 monomer
15 comprised residues 3–347. Residues 3–4, 30–44, 231–236, 333–335, and 341–346 of
16 the apo-AKR2E5 exhibited weak density and were not modeled. In the NADPH-
17 AKR2E5 structure, residues 231–236 were modeled. The occupancy for NADPH in the
18 refinement was 0.91. The structure of the NADPH-AKR2E5 complex indicated that
19 NADPH interacted with the following AKR2E5 residues: Trp25, Trp138, Ser175,
20 Asn176, Gln197, Ser224, Phe226, Ser232, Gly227, Lys277, Thr278, Val279, Thr280,
21 Arg283, Gln286, and Asn287 (Supplementary Fig. 2). Arg283 in AKR2E5 interacted
22 with the 2'-phosphate group (O2X) of NADPH via hydrogen bonding, and the
23 pyrophosphate group of NADPH was also stabilized by Trp25 (via a water molecule),
24 Ser224, Phe226, Gly227 (via a water molecule), and Lys277. We found that Trp138

(via a water molecule), Ser175, Asn176, and Gln197 were involved in the interaction with the amide in the nicotinamide ring, which oriented the ring and controlled hydride transfer stereospecificity. The conformation of the nicotinamide ring was further constrained by stacking interactions with the Tyr223 side chain. In addition, NADPH 2'-phosphate group was bound to Lys277, Thr278, Val279, Thr280, and Arg283 via hydrogen bonds and charge complementarity with positively charged amino acid side chains. Finally, the adenosine part of NADPH formed hydrogen bonds with Ser232, Gln286 (via a water molecule), and Asn287.

AKR2E5 gel-filtration profile indicated that the enzyme was a monomer.

AKR2E5 showed similarity to most AKRs, although dimers [17,18], trimers (unpublished, PDB ID: 1OG6), tetramers [14], and octamers [19] have also been observed. The overall AKR2E5 structure consisted of a TIM-barrel in the central part of the enzyme, which was also similar to other AKR superfamily members. Moreover, we observed an extended loop at the C-terminal region, which is important for efficient substrate recognition in other AKRs [20,21]. In the NADPH-AKR2E5 structure, residues 231–236 corresponding to the extended loop were modeled, whereas in the apo-AKR2E5 structure this loop was disordered, indicating that NADPH binding stabilized the C-terminal region of AKR2E5.

Previously, we determined the crystal structure of AKR2E4 (PDB ID: 3WCZ) [15]. Superimposition of AKR2E4 and AKR2E5 revealed that the structures of the two enzymes were similar with a root-mean-square deviation (C^α) of 0.992 Å/306 residues. The AKR2E5 structure had loop 1 (residues 30–44) and the C-terminal region (residues 333–347), which are absent in AKR2E4, suggesting that both regions could define the difference in substrate specificity between these AKRs.

3.3. *Substrate and cosubstrate specificity*

Recombinant AKR2E5 expressed in bacteria was soluble and purified to homogeneity. We obtained 4.5 mg of purified AKR2E5 from 1 L of medium. Kinetic parameters of the purified AKR2E5 for DL-glyceraldehyde (DLG), methylglyoxal, 17 α -hydroxyprogesterone, and 3-hydroxybenzaldehyde were determined using NADPH as a cosubstrate (Table 2).

AKR2E5 was unable to reduce methylglyoxal, a common substrate of AKRs (Table 2). Furthermore, unlike some members of AKR families 1, 2, 7, and 11, NADH was not an alternative cosubstrate for AKR2E5: no measurable AKR2E5 activity was detected with NADH for any of the substrates listed in Table 2. In the presence of NADPH, AKR2E5 exhibited substrate preferences for 11-hexadecenal and bombykal, two pheromonal components in insects. AKR2E5 recognized DLG and xylose as substrates; however, no measurable activity with glucose was detected.

3.4. *Amino acid residues responsible for catalysis*

The C-terminal region has been found to be important for efficient substrate binding in other AKRs [20,21], and AKR2E5 was shown to have a C-terminal extension, suggesting distinct substrate specificity.

The catalytic tetrad (Asp, Tyr, Lys, His) is crucial for AKR activity [22,23]. In this tetrad, Tyr functions as a proton donor for the substrate reduction by AKR, whereas Lys is charged by interacting with Asp, which facilitates Tyr protonation through a

hydrogen bond between Lys and Tyr, resulting in the protonation of the carbonyl substrate by proton relay. Finally, His participates in substrate orientation and stabilization. Based on the refined AKR2E4 structure, we have previously found that its catalytic residues (Asp53, Tyr58, Lys87, and His120) are strictly conserved across all AKRs [15]. Superimposition of AKR2E4 and AKR2E5 revealed that the corresponding residues in AKR2E5 were Asp70, Tyr75, Lys104, and His137. Fig. 2 shows AKR2E5-bound NADPH as a Fo-Fc electron density map. To verify that the four residues were involved in AKR2E5 activity, we substituted each amino acid with Ala by site-directed mutagenesis and measured specific kinetic activities of the generated mutants (D70A, Y75A, K104A, and H137A) with various substrates. Since the activity of AKR2E5 mutants did not fit the Michaelis-Menten equation, we determined specific activity with DL-glyceraldehyde (10 mM), 3-hydroxy benzaldehyde (10 mM), 17 α -hydroxy progesterone (10 mM), 11-hexadecenal (8.4 mM), and bombykal (1.7 mM) as substrates. The results indicate that the catalytic activity of AKR2E5 mutants was lower than that of the wild type (Fig. 3). The mutant enzymes did not show any activity toward xylose (data not shown).

The AKR family consists of aldose reductases and aldehyde reductases [3,24]. As shown in Table 1, AKR2E5 demonstrated both activities; however, AKR2E5 affinity to xylose was not high as evidenced by the K_m value which was about 3-fold higher than that of *Candida tenuis* xylose reductase [24]. According to k_{cat}/K_m , catalytic efficiency, AKR2E5 had the highest activity with glyceraldehyde, 11-hexadecenal, and bombykal, even though the K_m s values for glyceraldehyde and 11-hexadecenal were higher than that for bombykal. When 11-hexadecenal was used as a substrate, the D70A mutant had negligible activity, whereas the other three mutants demonstrated 20–40% activity of

the wild type. On the other hand, when bombykal was used as a substrate, the D70A mutant showed ~30% activity of the wild type and the other three mutants had no measurable activity. As the structures of these substrates are similar, the reason for differential substrate specificity is still unclear. Given that a mixture of sex pheromones bombykal ([10E,12Z]-hexadecadienal) and bombykol ([10E,12Z]-hexadecadienal) is produced by female silkworms [25] to stimulate males [26], a major function of AKR2E5 is likely the production of silkworm pheromones. Moreover, although the composition and ratios of pheromone blends exhibit interspecific variation (<http://www.pherobase.com>), the kinetic data on AKR2E5 indicate that it is also able to reduce 11-hexadecenal, one of the components in the *Helicoverpa armigera* pheromone blend (also containing 9-hexadecenal) [27]. Since 11-hexadecenal and bombykal share similar chemical structures, we suspect that AKR2E5 is able to recognize 11-hexadecenal as a substrate analogue in the catalytic reaction. Indeed, an AKR that uses NADPH as a cosubstrate to reduce 9-hexadecenal and 11-hexadecenal has been found in *H. armigera* [27]. Real-time quantitative PCR analysis revealed AKR2E5 mRNA in adult tissues including antenna, fat body, head, and pheromone gland (data not shown). As the pheromone gland and antenna are the sites of pheromone production and detection, respectively [26], these findings indicate that AKR2E5 functions as a bombykal reductase. The results of our study together with previous findings support the hypothesis that AKR2E5 influences bombykol metabolism in the silkworm. Future studies could focus on altering AKR2E5 expression in silkworms to examine its influence on mating behavior.

In conclusion, this study identified and biochemically characterized AKR2E5, a novel AKR in the silkworm *B. mori*. Similar to other AKRs, AKR2E5 reduces various

1 substrates. Moreover, we identified AKR2E5 amino acid residues that are likely to play
2 important roles in catalysis. Finally, kinetic and structural analyses indicate that
3 AKR2E5 could have a function in the pheromone metabolism of *B. mori*.

4 5 **Acknowledgments**

6
7 This work was supported by a Grant-in-Aid for Scientific Research (KAKENHI,
8 24580078) from the Ministry of Education, Culture, Sports, Science and Technology of
9 Japan, and by a Research Grant for Young Investigators of the Department of
10 Agriculture, Kyushu University. We are grateful to Professor David K. Wilson at the
11 University of California-Davis for useful discussions of the manuscript. This work was
12 performed using a synchrotron beamline BL44XU at SPring-8 under the Cooperative
13 Research Program of Institute for Protein Research, Osaka University. Diffraction data
14 were collected at the Osaka University beamline BL44XU at SPring-8 (Harima, Japan)
15 under proposal numbers 2014A6961, 2014B6961, and 2015A6555.

References

- [1] K.M. Bohren, B. Bullock, B. Wermuth, K.H. Gabbay, The aldo-keto reductase superfamily. cDNAs and deduced amino acid sequences of human aldehyde and aldose reductases, *J Biol Chem* 264 (1989) 9547-9551.
- [2] Y. Jin, T.M. Penning, Aldo-keto reductases and bioactivation/detoxication, *Annu Rev Pharmacol Toxicol* 47 (2007) 263-292.
- [3] T.M. Penning, J.E. Drury, Human aldo-keto reductases: Function, gene regulation, and single nucleotide polymorphisms, *Arch Biochem Biophys* 464 (2007) 241-250.
- [4] K. Mita, M. Kasahara, S. Sasaki, Y. Nagayasu, T. Yamada, H. Kanamori, N. Namiki, M. Kitagawa, H. Yamashita, Y. Yasukochi, K. Kadono-Okuda, K. Yamamoto, M. Ajimura, G. Ravikumar, M. Shimomura, Y. Nagamura, I.T. Shin, H. Abe, T. Shimada, S. Morishita, T. Sasaki, The genome sequence of silkworm, *Bombyx mori*, *DNA Res* 11 (2004) 27-35.
- [5] K. Mita, M. Morimyo, K. Okano, Y. Koike, J. Nohata, H. Kawasaki, K. Kadono-Okuda, K. Yamamoto, M.G. Suzuki, T. Shimada, M.R. Goldsmith, S. Maeda, The construction of an EST database for *Bombyx mori* and its application, *Proc Natl Acad Sci U S A* 100 (2003) 14121-14126.
- [6] U.K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature* 227 (1970) 680-685.
- [7] W. Kabsch, Xds, *Acta Crystallogr D Biol Crystallogr* 66 (2010) 125-132.
- [8] P.R. Evans, An introduction to data reduction: space-group determination, scaling and intensity statistics, *Acta Crystallogr D Biol Crystallogr* 67 (2011) 282-292.
- [9] A.J. McCoy, R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni, R.J.

- Read, Phaser crystallographic software, *J Appl Crystallogr* 40 (2007) 658-674.
- [10] F. Faucher, K. Pereira de Jesus-Tran, L. Cantin, V. Luu-The, F. Labrie, R. Breton, Crystal structures of mouse 17alpha-hydroxysteroid dehydrogenase (apoenzyme and enzyme-NADP(H) binary complex): identification of molecular determinants responsible for the unique 17alpha-reductive activity of this enzyme, *J Mol Biol* 364 (2006) 747-763.
- [11] P.D. Adams, P.V. Afonine, G. Bunkoczi, V.B. Chen, N. Echols, J.J. Headd, L.W. Hung, S. Jain, G.J. Kapral, R.W. Grosse Kunstleve, A.J. McCoy, N.W. Moriarty, R.D. Oeffner, R.J. Read, D.C. Richardson, J.S. Richardson, T.C. Terwilliger, P.H. Zwart, The Phenix software for automated determination of macromolecular structures, *Methods* 55 (2011) 94-106.
- [12] P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics, *Acta Crystallogr D Biol Crystallogr* 60 (2004) 2126-2132.
- [13] V.B. Chen, W.B. Arendall, 3rd, J.J. Headd, D.A. Keedy, R.M. Immormino, G.J. Kapral, L.W. Murray, J.S. Richardson, D.C. Richardson, MolProbity: all-atom structure validation for macromolecular crystallography, *Acta Crystallogr D Biol Crystallogr* 66 (2010) 12-21.
- [14] H. Oritani, Y. Deyashiki, T. Nakayama, A. Hara, H. Sawada, K. Matsuura, Y. Bunai, I. Ohya, Purification and characterization of pig lung carbonyl reductase, *Arch Biochem Biophys* 292 (1992) 539-547.
- [15] K. Yamamoto, D.K. Wilson, Identification, characterization, and crystal structure of an aldo-keto reductase (AKR2E4) from the silkworm *Bombyx mori*, *Arch Biochem Biophys* 538 (2013) 156-163.
- [16] W. Kabsch, C. Sander, Dictionary of protein secondary structure: pattern

1 recognition of hydrogen-bonded and geometrical features, Biopolymers 22
2 (1983) 2577-2637.

3 [17] K.L. Kavanagh, M. Klimacek, B. Nidetzky, D.K. Wilson, The structure of apo and
4 holo forms of xylose reductase, a dimeric aldo-keto reductase from *Candida*
5 *tenuis*, Biochemistry 41 (2002) 8785-8795.

6 [18] E. Kozma, E. Brown, E.M. Ellis, A.J. Lapthorn, The crystal structure of rat liver
7 AKR7A1. A dimeric member of the aldo-keto reductase superfamily, J Biol
8 Chem 277 (2002) 16285-16293.

9 [19] A.J. Lapthorn, X. Zhu, E.M. Ellis, The diversity of microbial aldo/keto reductases
10 from *Escherichia coli* K12, Chem Biol Interact 202 (2013) 168-177.

11 [20] K.M. Bohren, C.E. Grimshaw, K.H. Gabbay, Catalytic effectiveness of human
12 aldose reductase. Critical role of C-terminal domain, J Biol Chem 267 (1992)
13 20965-20970.

14 [21] K.J. Rees-Milton, Z. Jia, N.C. Green, M. Bhatia, O. El-Kabbani, T.G. Flynn,
15 Aldehyde reductase: the role of C-terminal residues in defining substrate and
16 cofactor specificities, Arch Biochem Biophys 355 (1998) 137-144.

17 [22] J.M. Jez, M.J. Bennett, B.P. Schlegel, M. Lewis, T.M. Penning, Comparative
18 anatomy of the aldo-keto reductase superfamily, Biochem J 326 (Pt 3) (1997)
19 625-636.

20 [23] D.K. Wilson, K.M. Bohren, K.H. Gabbay, F.A. Quiocho, An unlikely sugar
21 substrate site in the 1.65 Å structure of the human aldose reductase holoenzyme
22 implicated in diabetic complications, Science 257 (1992) 81-84.

23 [24] R. Kratzer, S. Leitgeb, D.K. Wilson, B. Nidetzky, Probing the substrate binding site
24 of *Candida tenuis* xylose reductase (AKR2B5) with site-directed mutagenesis,

Biochem J 393 (2006) 51-58.

[25] K.E. Kaissling, G. Kasang, H.J. Bestmann, W. Stransky, O. Vostrowsky, A New

Pheromone of the Silkworm Moth *Bombyx mori*.

Sensory pathway and behavioral effect, *Naturwissenschaften* 65 (1978) 382–384.

[26] J.A. Tillman, S.J. Seybold, R.A. Jurenka, G.J. Blomquist, Insect pheromones--an

overview of biosynthesis and endocrine regulation, *Insect Biochem Mol Biol* 29

(1999) 481-514.

[27] H. Guo, A. Del Corso, L.Q. Huang, U. Mura, P. Pelosi, C.Z. Wang, Aldehyde

reductase activity in the antennae of *Helicoverpa armigera*, *Insect Mol Biol* 23

(2014) 330-340.

Figure Legends

Fig. 1. Tertiary structure of AKR2E5. Bound NADPH is shown in blue. The names of the helices and strands as they appear in Supplementary Figure 1 are adjacent to each structural element. N- and C-terminal amino acid residues are indicated as N and C, respectively. The figure was drawn using PyMOL (<http://pymol.sourceforge.net>).

Fig. 2. Structure of the catalytic tetrad. The σ_A -weighted F_O-F_C electron density map of NADPH was contoured at 1.5σ . The catalytic tetrad composed of Asp70, Tyr75, Lys104, and His137 is shown.

Fig. 3. Specific activity of AKR2E5 mutants with various substrates. The catalytic activity of the wild-type and mutant (D70A, Y75A, K104A, and H137A) AKR2E5 was measured with the substrates: DL-glyceraldehyde (10 mM), 3-hydroxy benzaldehyde (10 mM), 17α -hydroxy progesterone (10 mM), 11-hexadecenal (8.4 mM), and bombykal (1.7 mM). The data are presented as the mean $\pm$ SD of three independent experiments.

Tables

Table 1. Data collection and refinement statistics for AKR2E5 (Values in parentheses indicate the highest-resolution shell)

Parameter	apo-AKR2E5	AKR2E5 complexed with NADPH
Space group	<i>P</i> 6 ₁ 22	<i>P</i> 6 ₁ 22
Unit cell parameters (Å)	<i>a</i> = <i>b</i> = 112.18, <i>c</i> = 147.63	<i>a</i> = <i>b</i> = 111.95, <i>c</i> = 147.65
Beam line	SPring-8, BL44XU	SPring-8, BL44XU
Wavelength (Å)	0.9	0.9
Resolution range (Å)	37.0–2.20 (2.24–2.20)	46.1–2.30 (2.34–2.30)
Total number of reflections	343,695	383,297
Number of unique reflections	28,491 (1,398)	24,781 (1,215)
Multiplicity	12.1 (12.3)	15.5 (15.4)
<i>R</i> _{merge} ^a (%)	7.7 (48.8)	7.3 (46.3)
<i><I>/<σ(I)></i>	35.6 (6.97)	30.0 (5.04)
Completeness (%)	99.8 (99.7)	99.7 (99.3)
Refinement statistics		
Resolution range (Å)	37.0–2.20	46.1–2.30
Number of reflections	27,609	23,541
<i>R</i> _{work} ^b (%) / <i>R</i> _{free} ^c (%)	15.7/20.4	15.9/20.6
Root-mean-square deviations		
Bond lengths (Å)/Bond angles (°)	0.021/1.909	0.019/1.935
Average B-factors (Å ²)/Number of atoms		
Protein	34.8/2,614	35.4/2,637
Small molecules ^d	42.8/37	44.7/77
Water	41.3/252	41.0/241
Ramachandran plot		
Favored region (%)	96.9	96.3
Allowed regions (%)	3.1	3.7

^a $R_{\text{merge}} = \sum (I - \langle I \rangle) / \sum \langle I \rangle$, where I is the intensity measurement for a given reflection and $\langle I \rangle$ is the average intensity for multiple measurements of this reflection.

^b $R_{\text{work}} = \sum |F_{\text{obs}} - F_{\text{cal}}| / \sum F_{\text{obs}}$, where F_{obs} and F_{cal} are the observed and calculated structure-factor amplitudes.

^cThe R_{free} value was calculated using only an unrefined, randomly chosen subset of reflection data (5%) that were excluded from refinement.

^dSmall molecules include NADPH, ethylene glycol, acetate ion and calcium ion.

Table 2. Kinetic parameters for several substrates

Substrate	K_m (mM)	k_{cat} (min ⁻¹)	k_{cat} / K_m (min ⁻¹ mM ⁻¹)
DL-glyceraldehyde	20 (3.7)	7.6 (1.3)	0.38
Methylglyoxal	NMA	NMA	NMA
17 α -hydroxy progesterone	18 (2.2)	1.5 (0.21)	0.084
3-hydroxy benzaldehyde	11 (1.6)	0.52 (0.12)	0.046
11-hexadecenal	4.4 (1.3)	2.5 (0.48)	0.57
bombykal	0.080 (0.022)	0.030 (0.012)	0.38
xylose	300 (79)	0.25 (0.066)	8.4×10^{-4}
glucose	NMA	NMA	NMA

Data except k_{cat} / K_m are means (SD) of three independent experiments. NMA, no measurable activity.

Figure

Fig. 1

Fig. 2

Fig. 3

Fig. 3_continued